# HUMAN SARCOMA-ASSOCIATED ANTIGENS

## Related Applications

This application is a continuation-in-part of U.S. application serial no. 10/260,708, filed on September 30, 2002, now pending, the disclosure of which is incorporated herein by reference.

## Field of the Invention

The invention relates to sarcoma-associated antigens and the nucleic acid molecules that encode them. The invention further relates to the use of the nucleic acid molecules, polypeptides and fragments thereof associated with sarcoma in methods and compositions for the diagnosis and treatment of diseases, such as cancer. More specifically, the invention relates to the discovery of a novel cancer/testis (CT) antigen, NY-SAR-35.

## Background of the Invention

The identification of human tumor antigens recognized by the autologous host is yielding new and promising target molecules for immunotherapy, diagnosis and monitoring of human cancer (van der Bruggen P, et al. 1991. A gene encoding an antigen recognized by cytolytic T lymphocytes on a human melanoma. Science 254:1643-47; Gaugler, B., et al. Human gene MAGE-3 codes for an antigen recognized on a melanoma by autologous cytolytic T lymphocytes. *J. Exp. Med.* 1994; 179: 921-30; Kawakami, Y., et al. Cloning of the gene for a shared human melanoma antigen recognized by autologous T cells infiltrating into tumor. *Proc. Natl. Acad. Sci. USA.* 1994; 91: 3515-19 and Chen, Y. -T., et al. A testicular antigen aberrantly expressed in human cancers detected by autologous antibody screening. *Proc. Natl. Acad. Sci. USA.* 1997; 94: 1914-18). Studies of the cellular and humoral immune response to cancer have revealed an extensive repertoire of tumor antigens recognized by the immune system, collectively termed the cancer immunome (Jager D, et al. Identification of a tissue-specific putative transcription factor in breast tissue by serological screening of a breast cancer library. Cancer Res 2001 Mar 1;61(5):2055-61).

The immunome is composed largely of antigens defined by T-cell epitope cloning (van der Bruggen P, et al. 1991. A gene encoding an antigen recognized by cytolytic T lymphocytes on a human melanoma. Science 254:1643-47; Gaugler, B., et al. Human gene MAGE-3 codes for an antigen recognized on a melanoma by autologous cytolytic T

lymphocytes. *J. Exp. Med.* 1994; 179: 921-30; Kawakami, et al. Cloning of the gene for a shared human melanoma antigen recognized by autologous T cells infiltrating into tumor. *Proc. Natl. Acad. Sci. USA.* 1994; 91: 3515-19; Boel, P., et al. BAGE: a new gene encoding an antigen recognized on human melanomas by cytolytic T lymphocytes. *Immunity* 1995; *2:* 167-75. (PMID: 7895173); Van den Eynde, B., et al. A new family of genes coding for an antigen recognized by autologous cytolytic T lymphocytes on a human melanoma. *J. Exp. Med.* 1995; 182: 689-98. (PMID: 7544395)), MHC peptide elution (Skipper JC, et al. An HLA-A2-restricted tyrosinase antigen on melanoma cells results from posttranslational modification and suggests a novel pathway for processing of membrane proteins. J Exp Med 1996 Feb 1;183(2):527-34; Cox AL, et al. Identification of a peptide recognized by five melanoma-specific human cytotoxic T cell lines. Science 1994 Apr 29;264(5159):716-9; Pascolo S, et al. A MAGE-A1 HLA-A A*0201 epitope identified by mass spectrometry. Cancer Res 2001 May 15;61(10):4072-7), and serological expression cloning (SEREX, Chen, Y. -T., et al. A testicular antigen aberrantly expressed in human cancers detected by autologous antibody screening. *Proc. Natl. Acad. Sci. USA.* 1997; 94: 1914-18; Jager D, et al. Identification of a tissue-specific putative transcription factor in breast tissue by serological screening of a breast cancer library. Cancer Res 2001 Mar 1;61(5):2055-61; Sahin, U., et al. Human neoplasms elicit multiple specific immune responses in the autologous host. *Proc. Natl. Acad. Sci. USA* 1995; 92: 11810-13; Scanlan, M. J., et al. Characterization of human colon cancer antigens recognized by autologous antibodies. *Int. J. Cancer* 1998; 76: 652-8; Scanlan, M. J., et al. Antigens recognized by autologous antibody in patients with renal-cell carcinoma. *Int. J. Cancer* 1999; 83: 456-64; Scanlan MJ, et al. Humoral immunity to human breast cancer: antigen definition and quantitative analysis of mRNA expression. Cancer Immunity 1:4 [epub]), and is catalogued in three databases: the peptide database of T-cell defined tumor antigens (authored by members of the Ludwig Institute for Cancer Research (LICR) that is available on the website of *Cancer Immunity,* Journal of the Academy of Cancer Immunology, http://www.cancerimmunity.org/peptidedatabase/Tcellepitopes); the SYFPEITHI database of MHC ligands and peptide motifs (available on the website of Biomedical Informatics-Heidelberg, http://www.bmi-heidelberg.com/syfpeithi/) and the cancer immunome database available on the website of the LICR (www2.licr.org/CancerImmunomeDB, formerly www.licr.org/SEREX.html).

SEREX is a method of immunoscreening tumor-derived cDNA expression libraries with cancer patient sera in order to identify molecules recognized by high titered IgG

antibodies (Sahin, U., et al. Human neoplasms elicit multiple specific immune responses in the autologous host. *Proc. Natl. Acad. Sci. USA* 1995; 92: 11810-13) Approximately 1000 distinct antigens have been defined by SEREX analysis, including a number of etiologically and therapeutically significant cancer antigens, such as mutational antigens (e.g. p53, LKB1, BUB1; Scanlan, M. J., et al. Characterization of human colon cancer antigens recognized by autologous antibodies. *Int. J. Cancer* 1998; 76: 652-8; Scanlan, M. J., et al. Antigens recognized by autologous antibody in patients with renal-cell carcinoma. *Int. J. Cancer* 1999; 83: 456-64; Scanlan MJ, et al. Humoral immunity to human breast cancer: antigen definition and quantitative analysis of mRNA expression. Cancer Immunity 1:4 [epub]), differentiation antigens (e.g. tyrosinase, NY-BR-1, rab 38; Jager D, et al. Identification of a tissue-specific putative transcription factor in breast tissue by serological screening of a breast cancer library. Cancer Res 2001 Mar 1;61(5):2055-61; Sahin, U., et al. Human neoplasms elicit multiple specific immune responses in the autologous host. *Proc. Natl. Acad. Sci. USA* 1995; 92: 11810-13; Jager D, et al. Serological cloning of a melanocyte rab guanosine 5'-triphosphate-binding protein and a chromosome condensation protein from a melanoma complementary DNA library. Cancer Res 2000 Jul 1;60(13):3584-91), overexpressed gene products (e.g. Her2neu, TPD52, eIF4-gamma; Scanlan MJ, et al. Humoral immunity to human breast cancer: antigen definition and quantitative analysis of mRNA expression. Cancer Immunity 1:4 [epub]; Chen, Y. -T., et al. Identification of human tumor antigens by serological expression cloning. *In:* S. A. Rosenberg (ed.). Principles and Practice of Biologic Therapy of Cancer, pp. 557- 570. Philadelphia: Lippincott Williams & Wilkins, 2000) and cancer/testis (CT) antigens (e.g. MAGE-1, NY-ESO-1, SSX-2; Chen, Y. -T., et al. A testicular antigen aberrantly expressed in human cancers detected by autologous antibody screening. *Proc. Natl. Acad. Sci. USA.* 1997; 94: 1914-18; Sahin, U., et al. Human neoplasms elicit multiple specific immune responses in the autologous host. *Proc. Natl. Acad. Sci. USA* 1995; 92: 11810-13).

CT antigens represent a group of shared, tumor-specific antigens expressed exclusively in developing germ cells of the testis and fetal ovary, as well as in placental trophoblast, and most notably, in a proportion of human cancers of diverse origins (Chen, Y. -T., et al. Identification of human tumor antigens by serological expression cloning. *In:* S. A. Rosenberg (ed.). Principles and Practice of Biologic Therapy of Cancer, pp. 557- 570. Philadelphia: Lippincott Williams & Wilkins, 2000). These antigens elicit spontaneous cellular (Van den Eynde, B. J. and van der Bruggen, P. (1997) Curr. Opin. Immunol. 9,684-

693) and humoral immune responses (Stockert, E., et al. (1998) J. Exp. Med. 187, 1349-1354) in some cancer patients. On the basis of tissue-restricted expression and immunogenicity, CT antigens are attractive targets for vaccine-based immunotherapies. In general, CT antigens are expressed in 20-40% of specimens from a given tumor type (Sahin U, et al. 1998. Expression of multiple cancer/testis antigens in breast cancer and melanoma: basis for polyvalent CT vaccine strategies. Int J Cancer 78:387-89; Scanlan MJ et al. 2000. Expression of cancer-testis antigens in lung cancer: definition of bromodomain testis-specific gene (BRDT) as a new CT gene, CT9. Cancer Lett. 150:155-64; Van den Eynde BJ and van der Bruggen P. 1997. T cell defined tumor antigens. Curr Opin Immunol 9:684-693). One exception to this is synovial sarcoma, in which 80% of specimens express NY-ESO-1 (Jungbluth AA, et al. 2001. Monophasic and biphasic synovial sarcomas abundantly express cancer/testis antigen NY-ESO-1 but not MAGE-A1 or CT7. Int J Cancer 94:252-6) and MAGE antigens (Antonescu CR, et al. MAGE antigen expression in monophasic and biphasic synovial sarcoma. Hum Pathol 2002 Feb;33(2):225-9); the expression of which are often homogeneous throughout the tumor. Thus, identification of additional CT antigens and other genes having a tumor-associated expression profile is needed for the development of additional therapeutics and diagnostics to permit effective treatment and diagnosis of a broader group of cancer patients.

## Summary of the Invention

The humoral immune response of sarcoma patients to CT antigens was examined using the SEREX method. Sera from patients which showed a humoral immune response to CT antigens were subsequently used to screen cDNA libraries derived from CT-rich synovial sarcoma cell lines as well as normal testis. Although there was little overlap in the identity of clones isolated with different sarcoma sera, more than 30% of the isolated clones were previously identified during SEREX analysis of other tumor types. Approximately 60% of these antigens also reacted with sera from normal individuals. This is in conformity with other findings (Scanlan, M. J., et al. Antigens recognized by autologous antibody in patients with renal-cell carcinoma. *Int. J. Cancer* 1999; 83: 456-64 and Scanlan MJ, et al. Humoral immunity to human breast cancer: antigen definition and quantitative analysis of mRNA expression. Cancer Immunity 20001; 1:4 [epub]). Thus, only a fraction of the serologically-defined immunome is associated with a cancer-related immune response. The studies described herein have led to the identification of antigens, which include antigens not before

associated with cancer along with several novel gene products associated with a sarcoma-related immune response. One such novel CT antigen is NY-SAR-35, which appears to be a cell surface/secreted molecule.

According to one aspect of the invention, isolated nucleic acid molecules are provided. The isolated nucleic acid molecules are selected from the group consisting of (a) nucleic acid molecules which hybridize under high stringency conditions to a nucleic acid molecule comprising a nucleotide sequence selected from the group consisting of nucleotide sequences set forth as SEQ ID NOs: 1-14 and 97-107 and which code for a sarcoma-associated antigen, (b) nucleic acid molecules that differ from the nucleic acid molecules of (a) in codon sequence due to the degeneracy of the genetic code, and (c) complements of (a) or (b).

In some embodiments, the isolated nucleic acid molecule includes a nucleotide sequence selected from the group consisting of nucleotide sequences set forth as SEQ ID NOs: 1-14 and 97-107. In some embodiments the isolated nucleic acid molecule comprises a nucleotide sequence selected from the group consisting of nucleotide sequences set forth as SEQ ID NOs: 10, 11, 99, 102 and 104. In other embodiments the nucleic acid molecule comprises a nucleotide sequence set forth as SEQ ID NO: 10. In yet other embodiments the nucleic acid molecule comprises a nucleotide sequence set forth as SEQ ID NO: 11. In still other embodiments the nucleic acid molecule comprises a nucleotide sequence set forth as SEQ ID NO: 102. In still further embodiments the nucleic acid molecule comprises a nucleotide sequence set forth as SEQ ID NO: 104.

In some embodiments the nucleic acid molecule comprises a nucleotide sequence set forth as SEQ ID NOs: 121, 123, 125, 127, 129 or 131. In some embodiments the nucleic acid molecule comprises a nucleotide sequence set forth as SEQ ID NO: 121. In other embodiments the nucleic acid molecule comprises a nucleotide sequence set forth as SEQ ID NO: 123. In still other embodiments the nucleic acid molecule comprises a nucleotide sequence set forth as SEQ ID NO: 125. In yet other embodiments the nucleic acid molecule comprises a nucleotide sequence set forth as SEQ ID NO: 131.

According to another aspect of the invention, additional isolated nucleic acid molecules are provided. The isolated nucleic acid molecules are selected from the group consisting of: (a) unique fragments of a nucleotide sequence selected from the group consisting of nucleotide sequences set forth as SEQ ID NOs: 10, 11, 99, 102 and 104, which encodes an immunogenic peptide and (b) complements of (a). In some embodiments the

isolated nucleic acid molecules are selected from the group consisting of: (a) unique fragments of a nucleotide sequence set forth as SEQ ID NO: 10, which encodes an immunogenic peptide and (b) complements of (a). In other embodiments the isolated nucleic acid molecules are selected from the group consisting of: (a) unique fragments of a nucleotide sequence set forth as SEQ ID NO: 11, which encodes an immunogenic peptide and (b) complements of (a). In yet other embodiments the isolated nucleic acid molecules are selected from the group consisting of: (a) unique fragments of a nucleotide sequence set forth as SEQ ID NO: 102, which encodes an immunogenic peptide and (b) complements of (a). In still other embodiments the isolated nucleic acid molecules are selected from the group consisting of: (a) unique fragments of a nucleotide sequence set forth as SEQ ID NO: 104, which encodes an immunogenic peptide and (b) complements of (a). In some embodiments the nucleic acid molecule comprises a nucleotide sequence set forth as SEQ ID NO: 121. In other embodiments the nucleic acid molecule comprises a nucleotide sequence set forth as SEQ ID NO: 123. In still other embodiments the nucleic acid molecule comprises a nucleotide sequence set forth as SEQ ID NO: 125. In yet other embodiments the nucleic acid molecule comprises a nucleotide sequence set forth as SEQ ID NO: 131.

In certain embodiments, the isolated nucleic acid molecule includes a nucleotide sequence that is at least about 90% identical to a nucleotide sequence selected from the group consisting of SEQ ID NOs: 1-14 and 97-107; preferably the nucleotide sequence is at least about 95% identical, more preferably the nucleotide sequence is at least about 97% identical, still more preferably the nucleotide sequence is at least about 98% identical, and yet more preferably the nucleotide sequence is at least about 99% identical.

According to further aspects of the invention, expression vectors that include any of the foregoing isolated nucleic acid molecules operably linked to a promoter are provided, as are host cells transformed or transfected with these expression vectors. In certain embodiments, the host cell expresses a MHC molecule, and in some of these embodiments the MHC molecule is expressed recombinantly.

According to another aspect of the invention, isolated polypeptides are provided that are encoded by the isolated nucleic acid molecules described herein. In certain embodiments, the isolated polypeptide includes an amino acid sequence as set forth in SEQ ID NOs: 46-60, 109-120 or a fragment thereof that is at least eight amino acids in length. In certain embodiments, the isolated polypeptides are antigenic polypeptides that are capable of eliciting antibodies to a sarcoma-associated antigen. In some embodiments the isolated

polypeptide includes an amino acid sequence as set forth in SEQ ID NO: 55 or a fragment thereof that is at least eight amino acids in length. In other embodiments the isolated polypeptide includes an amino acid sequence as set forth in SEQ ID NO: 56 or a fragment thereof that is at least eight amino acids in length. In yet other embodiments the isolated polypeptide includes an amino acid sequence as set forth in SEQ ID NO: 111 or a fragment thereof that is at least eight amino acids in length. In still other embodiments the isolated polypeptide includes an amino acid sequence as set forth in SEQ ID NO: 114 or a fragment thereof that is at least eight amino acids in length. In yet other embodiments the isolated polypeptide includes an amino acid sequence as set forth in SEQ ID NO: 116 or a fragment thereof that is at least eight amino acids in length. In still other embodiments the isolated polypeptide includes an amino acid sequence as set forth in SEQ ID NO: 122. In yet other embodiments the isolated polypeptide includes an amino acid sequence as set forth in SEQ ID NO: 124. In still further embodiments the isolated polypeptide includes an amino acid sequence as set forth in SEQ ID NO: 126. In yet other embodiments the polypeptide includes the amino acid sequence set forth as SEQ ID NOs: 128, 130 or 132.

Another aspect of the invention provides binding polypeptides that selectively bind to the foregoing isolated polypeptides. In some embodiments these binding polypeptides are isolated also. In other embodiments, the binding polypeptides are antibodies or antigen-binding fragments thereof.

According to another aspect of the invention, methods of diagnosing cancer in a subject are provided. The methods include obtaining a biological sample from the subject, and determining the presence of an antibody in the biological sample that binds specifically to one or more sarcoma-associated antigens encoded by a nucleotide sequence selected from the group consisting of SEQ ID NOs: 3, 5-8, 10-45, 99, 102, 104 and 108. The presence of such antibodies indicates that the subject has cancer. In some embodiments the one or more sarcoma-associated antigens is/are encoded by a nucleotide sequence selected from the group consisting of SEQ ID NOs: 5-7, 10-13, 15-45, 102, 104 and 108. In still other embodiments the one or more sarcoma-associated antigens is/are encoded by a nucleotide sequence selected from the group consisting of SEQ ID NOs: 10, 11, 15, 102, 104 and 108. In some embodiments the sarcoma-associated antigens is encoded by a nucleotide sequence set forth as SEQ ID NO: 10.

In some embodiments, the step of determining the presence of an antibody includes contacting the biological sample with one or more sarcoma-associated antigens that are

specifically bound by the antibody and are encoded by a nucleic acid molecule comprising a nucleotide sequence selected from the group consisting of (1) nucleotide sequences set forth as SEQ ID NOs: 3, 5-8, 10-45, 99, 102, 104 and 108 and (2) nucleotide sequences that are at least 90% identical to the nucleotide sequences of (1), and then determining the binding of the antibody to the sarcoma-associated antigen. In other embodiments, the step of determining the presence of an antibody includes contacting the biological sample with one or more sarcoma-associated antigens that are specifically bound by the antibody and are encoded by a nucleic acid molecule comprising a nucleotide sequence selected from the group consisting of (1) nucleotide sequences set forth as SEQ ID NOs: 5-7, 10-13, 15-45, 102, 104 and 108 and (2) nucleotide sequences that are at least 90% identical to the nucleotide sequences of (1), and then determining the binding of the antibody to the sarcoma-associated antigen.

In some embodiments, the nucleic acid molecule comprises a nucleotide sequence selected from the group consisting of nucleotide sequences set forth as SEQ ID NOs: 10, 11, 15, 102, 104 and 108, and in other embodiments the nucleic acid molecule includes the nucleotide sequence set forth as SEQ ID NO: 10. In other embodiments the nucleic acid molecule includes the nucleotide sequence set forth as SEQ ID NO: 121. In still other embodiments the nucleic acid molecule includes the nucleotide sequence set forth in SEQ ID NO: 123. In other embodiments the nucleic acid molecule includes the nucleotide sequence set forth in SEQ ID NO: 125. In yet other embodiments the nucleic acid molecule includes the nucleotide sequence set forth in SEQ ID NOs: 127, 129 or 131.

In other embodiments, the sarcoma-associated antigen is a polypeptide that includes the amino acid sequence of any of SEQ ID NOs: 48, 50-53, 55-90, 111, 114, 116 and 120 or a fragment thereof that is at least eight amino acids in length. In still other embodiments, the sarcoma-associated antigen is a polypeptide that includes the amino acid sequence of any of SEQ ID NOs: 50-52, 55-58, 60-90, 114, 116, 120 or a fragment thereof that is at least eight amino acids in length. In still other embodiments, the sarcoma-associated antigen is a polypeptide that includes the amino acid sequence of any of SEQ ID NOs: 55, 56, 60, 114, 116 and 120 or a fragment thereof that is at least eight amino acids in length.

In some embodiments, the sarcoma-associated antigen includes the amino acid sequence set forth as SEQ ID NO: 55 or a fragment thereof that is at least eight amino acids in length. In certain embodiments the sarcoma-associated antigen includes the amino acid sequence set forth as SEQ ID NO: 122. In still other embodiments the sarcoma-associated

antigen includes the amino acid sequence set forth as SEQ ID NO: 124. In yet other embodiments the sarcoma-associated antigen includes the amino acid sequence set forth as SEQ ID NO: 126. In still other embodiments the sarcoma-associated antigen includes the amino acid sequence set forth as SEQ ID NO: 128, 130 or 132.

In certain embodiments, the biological sample is serum. In other embodiments, the one or more sarcoma-associated antigens are produced recombinantly, and/or the one or more sarcoma-associated antigens are bound to a substrate. In some embodiments, the step of determining the binding of the antibody with the one or more sarcoma-associated antigens is performed with an ELISA-based method. In still other embodiments a serum antibody detection assay (SADA) is used.

According to still another aspect of the invention, methods for diagnosing cancer in a subject are provided. The methods include obtaining a biological sample from a subject, and determining the expression of a sarcoma-associated antigen or a nucleic acid molecule that encodes it. The nucleic acid molecule includes a nucleotide sequence selected from the group consisting of SEQ ID NOs: 3, 5-8, 10-45, 99, 102, 104 and 108 in the biological sample. The nucleic acid molecule is some embodiments includes a nucleotide sequence selected from the group consisting of SEQ ID NOs: 5-7, 10-13, 15-45, 102, 104 and 108 in the biological sample. The expression of the sarcoma-associated antigen or the nucleic acid molecule that encodes it in the sample is diagnostic for cancer in the subject.

In certain embodiments, the sarcoma-associated nucleic acid molecule comprises the nucleotide sequence selected from the group consisting of SEQ ID NOs: 10, 11, 15, 102, 104 and 108. In some embodiments the sarcoma-associated nucleic acid molecule includes the nucleotide sequence set forth as SEQ ID NO: 10. In other embodiments the nucleic acid molecule includes the nucleotide sequence set forth as SEQ ID NO: 121. In still other embodiments the nucleic acid molecule includes the nucleotide sequence set forth in SEQ ID NO: 123. In yet other embodiments the nucleic acid molecule includes the nucleotide sequence set forth in SEQ ID NO: 125. In yet other embodiments the nucleic acid molecule includes the nucleotide sequence set forth in SEQ ID NOs: 127, 129 or 131.

In other embodiments, the sarcoma-associated antigen comprises an amino acid sequence selected from the group consisting of SEQ ID NOs: 48, 50-53, 55-90, 111, 114, 116 and 120 or a fragment thereof that is at least eight amino acids in length. In yet other embodiments, the sarcoma-associated antigen comprises an amino acid sequence selected from the group consisting of SEQ ID NOs: 50-52, 55-58, 60-90, 114, 116 and 120 or a

fragment thereof that is at least eight amino acids in length. The sarcoma-associated antigen in some embodiments includes an amino acid sequence selected from the group consisting of SEQ ID NOs: 55, 56, 60, 114, 116 and 120. In some embodiments the sarcoma-associated antigen includes an amino acid sequence set forth as SEQ ID NO: 55 or a fragment thereof that is at least eight amino acids in length. In certain embodiments the sarcoma-associated antigen includes the amino acid sequence set forth as SEQ ID NO: 122. In still other embodiments the sarcoma-associated antigen includes the amino acid sequence set forth as SEQ ID NO: 124. In yet other embodiments the sarcoma-associated antigen includes the amino acid sequence set forth as SEQ ID NO: 126. In still further embodiments the sarcoma-associated antigen includes the amino acid sequence set forth as SEQ ID NOs: 128, 130 or 132.

According to yet another aspect of the invention, methods for determining onset, progression, or regression of cancer in a subject are provided. The methods include obtaining from a subject a first biological sample, determining the expression of a sarcoma-associated antigen or the nucleic acid molecule that encodes it in the first sample, obtaining from the subject a second biological sample, determining the expression of the sarcoma-associated antigen or the nucleic acid molecule that encodes it in the second sample, and comparing the expression in the first sample to the expression in the second sample as a determination of the onset, progression, or regression of the cancer. The nucleic acid molecule comprises a nucleotide sequence selected from the group consisting of (1) nucleotide sequences set forth as SEQ ID NOs: 3, 5-8, 10-45, 99, 102, 104 and 108 and (2) nucleotide sequences that are at least 90% identical to the nucleotide sequences of (1). In some embodiments the nucleic acid molecule comprises a nucleotide sequence selected from the group consisting of (1) nucleotide sequences set forth as SEQ ID NOs: 5-7, 10-13, 15-45, 102, 104 and 108 and (2) nucleotide sequences that are at least 90% identical to the nucleotide sequences of (1).

In some embodiments, the nucleic acid molecule that encodes the sarcoma-associated antigen includes a nucleotide sequence selected from the group consisting of SEQ ID NOs: 10, 11, 15, 102, 104 and 108. In other embodiments the nucleic acid molecule includes the nucleotide sequence of SEQ ID NO: 10. In yet other embodiments the nucleic acid molecule includes the nucleotide sequence set forth as SEQ ID NO: 121. In still other embodiments the nucleic acid molecule includes the nucleotide sequence set forth in SEQ ID NO: 123. In yet other embodiments the nucleic acid molecule includes the nucleotide sequence set forth in SEQ ID NO: 125. In still other embodiments the nucleic acid molecule includes the

nucleotide sequence set forth in SEQ ID NOs: 127, 129 or 131. In other embodiments, the sarcoma-associated antigen includes a polypeptide sequence selected from the group consisting of polypeptide sequences set forth as SEQ ID NOs: 48, 50-53, 55-90, 111, 114, 116 and 120 or a fragment thereof that is at least eight amino acids in length. In still other embodiments, the sarcoma-associated antigen includes a polypeptide sequence selected from the group consisting of polypeptide sequences set forth as SEQ ID NOs: 50-52, 55-58, 60-90, 114, 116 and 120 or a fragment thereof that is at least eight amino acids in length. In yet other embodiments, the sarcoma-associated antigen includes a polypeptide sequence selected from the group consisting of polypeptide sequences set forth as SEQ ID NOs: 55, 56, 60, 114, 116 and 120 or a fragment thereof that is at least eight amino acids in length. In some embodiments the sarcoma-associated antigen includes the amino acid sequence of SEQ ID NO: 55 or a fragment thereof that is at least eight amino acids in length. In certain embodiments the sarcoma-associated antigen includes the amino acid sequence set forth as SEQ ID NO: 122. In still other embodiments the sarcoma-associated antigen includes the amino acid sequence set forth as SEQ ID NO: 124. In yet other embodiments the sarcoma-associated antigen includes the amino acid sequence set forth as SEQ ID NO: 126. In still other embodiments the sarcoma-associated antigen includes the amino acid sequence set forth as SEQ ID NOs: 128, 130 or 132.

In some embodiments of the foregoing methods, the step of determining the expression of the sarcoma-associated antigen or the nucleic acid molecule that encodes it includes contacting the biological sample with an agent that selectively binds to the sarcoma-associated antigen or the nucleic acid molecule that encodes it. For methods in which the agent that selectively binds is a nucleic acid molecule, it is preferred that the expression of the sarcoma-associated nucleic acid molecule is determined by nucleic acid hybridization or nucleic acid amplification; some embodiments of the methods utilize real-time RT-PCR or RT-PCR as methods of nucleic acid amplification, or use a nucleic acid microarray as a method for nucleic acid hybridization. For methods in which the agent that selectively binds is a polypeptide, the polypeptide preferably is an antibody or antigen-binding fragment thereof. More preferably, the antibody is a monoclonal antibody, particularly a chimeric, human, or humanized antibody, a single chain antibody, or the antigen-binding fragment is a $F(ab')_2$, Fab, Fd, or Fv fragment. In certain embodiments, the antibody or antigen-binding fragment is labeled with a detectable label, preferably a fluorescent or radioactive label.

In certain embodiments of the foregoing methods, the sample is selected from the group consisting of tissue, cells, and blood. In some embodiments, the cancer is a sarcoma.

In another aspect of the invention, kits for detecting antibodies reactive to a sarcoma-associated antigen in a biological sample are provided. The kits include one or more sarcoma-associated antigens encoded by a nucleic acid molecule comprising a nucleotide sequence selected from the group consisting of nucleotide sequences set forth as SEQ ID NOs: 3, 5-8, 10-45, 99, 102, 104 and 108, and instructions for the use of the sarcoma-associated antigens in the detection of antibodies in the biological sample. In some embodiments the one or more sarcoma-associated antigens is/are encoded by a nucleic acid molecule comprising a nucleotide sequence selected from the group consisting of nucleotide sequences set forth as SEQ ID NOs: 5-7, 10-13, 15-45, 102, 104 and 108. In some embodiments, the sarcoma-associated nucleic acid molecule comprises the nucleotide sequence set forth as SEQ ID NO: 10, 11, 15, 102, 104 or 108. In other embodiments, the sarcoma-associated nucleic acid molecule comprises the nucleotide sequence set forth as SEQ ID NO: 10. In other embodiments the nucleic acid molecule includes the nucleotide sequence set forth as SEQ ID NO: 121. In still other embodiments the nucleic acid molecule includes the nucleotide sequence set forth in SEQ ID NO: 123. In yet other embodiments the nucleic acid molecule includes the nucleotide sequence set forth in SEQ ID NO: 125. In still further embodiments the nucleic acid molecule includes the nucleotide sequence set forth in SEQ ID NOs: 127, 129 or 131. In other embodiments, the sarcoma-associated antigens are bound to a substrate. In further embodiments, the kit also includes a labeling reagent and labeling reagent substrate, and/or a blocking reagent. Additional kit embodiments include secondary antibodies for detection of the antibody bound to the antigen.

In a further aspect of the invention, other kits for the diagnosis of cancer in a subject are provided. The kits include one or more binding agents that specifically bind to a sarcoma-associated antigen or the nucleic acid molecule that encodes it. In this aspect, the nucleic acid molecule includes a nucleotide sequence selected from the group consisting of SEQ ID NOs: 3, 5-8, 10-45, 99, 102, 104 and 108. In some embodiments the nucleic acid molecule includes a nucleotide sequence selected from the group consisting of SEQ ID NOs: 5-7, 10-13, 15-45, 102, 104 and 108. The kit also includes instructions for the use of the binding agents in the diagnosis of cancer. The one or more binding agents are nucleic acid molecules or polypeptides. If the latter, the polypeptides preferably are antibodies or antigen-binding fragments thereof. In other embodiments, the one or more agents are bound to a

substrate. Further embodiments of the kits include one or more agents that bind specifically to a cancer-associated antigen other than those encoded by a nucleic acid molecule comprising a nucleotide sequence selected from the group consisting of SEQ ID NOs: 10, 11, 15, 102, 104 and 108. In some embodiments, the kit is configured for diagnosis of sarcomas.

According to another aspect of the invention, methods for treating a subject with a disorder characterized by the aberrant expression of a sarcoma-associated antigen or the nucleic acid molecule that encodes it are provided. The methods include administering to a subject an effective amount of an antibody or antigen-binding fragment thereof that specifically binds to the sarcoma-associated antigen. In this aspect, the antigen includes an amino acid sequence selected from the group consisting of SEQ ID NOs: 48, 50-53, 55-90, 111, 114, 116 and 120 or a fragment thereof that is eight or more amino acids in length. In some embodiments the antigen includes an amino acid sequence selected from the group consisting of SEQ ID NOs: 50-52, 55-58, 60-90, 114, 116 and 120 or a fragment thereof that is eight or more amino acids in length. In other embodiments the antigen includes an amino acid sequence selected from the group consisting of SEQ ID NOs: 55, 56, 60, 114, 116 and 120 or a fragment thereof that is eight or more amino acids in length. In some embodiments, the antibody or antigen-binding fragment thereof specifically binds to the extracellular domain of a sarcoma-associated antigen that includes the amino acid sequence of SEQ ID NO: 55 or a fragment thereof that is eight or more amino acids in length. In certain embodiments the sarcoma-associated antigen includes the amino acid sequence set forth as SEQ ID NO: 122. In still other embodiments the sarcoma-associated antigen includes the amino acid sequence set forth as SEQ ID NO: 124. In yet other embodiments the sarcoma-associated antigen includes the amino acid sequence set forth as SEQ ID NO: 126. In still other embodiments the sarcoma-associated antigen includes the amino acid sequence set forth as SEQ ID NO: 128, 130 or 132. In yet other embodiments the sarcoma-associated antigen includes the amino acid sequence set forth as SEQ ID NO: 134 or a fragment thereof that is eight or more amino acids in length.

In certain embodiments, the disorder is cancer, preferably sarcoma. In other embodiments, the antibody used in the methods is a monoclonal antibody, preferably a chimeric, human, or humanized antibody; a single chain antibody; or the antigen-binding fragment is a $F(ab')_2$, Fab, Fd, or Fv fragment.

In other embodiments, the antibody or antigen-binding fragment thereof is bound to a cytotoxic agent. Preferred cytotoxic agents include: calicheamicin, esperamicin,

methotrexate, doxorubicin, melphalan, chlorambucil, ARA-C, vindesine, mitomycin C, cisplatinum, etopside, bleomycin and 5-fluorouracil. Other cytotoxic agents include radioisotopes, including those that emit $\alpha$, $\beta$, and/or $\gamma$ radiation. Preferred radioisotopes include: $^{225}Ac$, $^{211}At$, $^{212}Bi$, $^{213}Bi$, $^{186}Rh$, $^{188}Rh$, $^{177}Lu$, $^{90}Y$, $^{131}I$, $^{67}Cu$, $^{125}I$, $^{123}I$, $^{77}Br$, $^{153}Sm$, $^{166}Bo$, $^{64}Cu$, $^{212}Pb$, $^{224}Ra$ and $^{223}Ra$.

According to another aspect of the invention, methods for treating a subject with a disorder characterized by the aberrant expression of a sarcoma-associated antigen or a nucleic acid molecule that encodes it are provided. The methods include administering an amount of an agent that selectively binds to the sarcoma-associated antigen or the nucleic acid molecule that encodes it effective to treat the disorder. The nucleic acid molecule comprises a nucleotide sequence selected from the group consisting of (a) an isolated nucleic acid molecule comprising a nucleotide sequence that is at least 90% identical to the nucleotide sequence selected from the group consisting of SEQ ID NOs: 3, 5-8, 10-45, 99, 102, 104 and 108, and (b) nucleic acid molecules that differ from the nucleic acid molecules of (a) in codon sequence due to the degeneracy of the genetic code. In some embodiments the nucleic acid molecule comprises a nucleotide sequence selected from the group consisting of (a) an isolated nucleic acid molecule comprising a nucleotide sequence that is at least 90% identical to the nucleotide sequence selected from the group consisting of SEQ ID NOs: 5-7, 10-13, 15-45, 102, 104 and 108, and (b) nucleic acid molecules that differ from the nucleic acid molecules of (a) in codon sequence due to the degeneracy of the genetic code. In certain embodiments the disorder is cancer, preferably sarcoma. In yet other embodiments the sarcoma-associated nucleic acid molecule comprises the nucleotide sequence selected from the group consisting of nucleotide sequences set forth as SEQ ID NOs: 10, 11, 15, 102, 104 and 108. In some embodiments the sarcoma-associated nucleic acid molecule comprises the nucleotide sequence set forth as SEQ ID NO: 10.

In other embodiments the sarcoma-associated nucleic acid molecule codes for a sarcoma-associated antigen which comprises the polypeptide sequence selected from the group consisting of polypeptide sequences set forth as SEQ ID NOs: 48, 50-53, 55-90, 111, 114, 116 and 120 or a fragment thereof that is at least eight amino acids in length. In still other embodiments the sarcoma-associated nucleic acid molecule codes for a sarcoma-associated antigen which comprises the polypeptide sequence selected from the group consisting of polypeptide sequences set forth as SEQ ID NOs: 50-52, 55-58, 60-90, 114, 116 and 120 or a fragment thereof that is at least eight amino acids in length. In some

embodiments the sarcoma-associated nucleic acid molecule codes for a sarcoma-associated antigen which comprises the polypeptide sequence set forth as SEQ ID NO: 55, 56, 60, 114, 116 or 120 or a fragment thereof that is at least eight amino acids in length. In another embodiment the sarcoma-associated nucleic acid molecule codes for a sarcoma-associated antigen which comprises the polypeptide sequence set forth as SEQ ID NO: 55 or a fragment thereof that is at least eight amino acids in length. In certain embodiments the sarcoma-associated antigen includes the amino acid sequence set forth as SEQ ID NO: 122. In still other embodiments the sarcoma-associated antigen includes the amino acid sequence set forth as SEQ ID NO: 124. In yet other embodiments the sarcoma-associated antigen includes the amino acid sequence set forth as SEQ ID NO: 126. In still other embodiments the sarcoma-associated antigen includes the amino acid sequence set forth as SEQ ID NOs: 128, 130 or 132.

In certain embodiments, the binding agent is an antisense or RNAi molecule. In other embodiments, the binding agent is a polypeptide, preferably an antibody or antigen-binding fragment thereof. Preferred antibodies include monoclonal antibodies, including chimeric, human, or humanized antibodies, and single chain antibodies; preferred antigen-binding fragments include $F(ab')_2$, Fab, Fd, or Fv fragments. In other embodiments, the antibody or antigen-binding fragment is bound to a cytotoxic agent.

According to yet another aspect of the invention, methods for treating a subject with a disorder characterized by the aberrant expression of a sarcoma-associated antigen or the nucleic acid molecule that encodes it are provided. The methods include administering to the subject an amount of an agent effective to stimulate an immune response to a sarcoma-associated antigen encoded by a nucleic acid molecule comprising a nucleotide sequence that is at least 90% identical to the nucleotide sequence selected from the group consisting of SEQ ID NOs: 3, 5-8, 10-45, 99, 102, 104 and 108. In some embodiments the sarcoma-associated antigen is encoded by a nucleic acid molecule comprising a nucleotide sequence that is at least 90% identical to the nucleotide sequence selected from the group consisting of SEQ ID NOs: 5-7, 10-13, 15-45, 102, 104 and 108. In some embodiments, the disorder is cancer, particularly sarcoma. In other embodiments the sarcoma-associated nucleic acid molecule comprises the nucleotide sequence selected from the group consisting of nucleotide sequences set forth as SEQ ID NOs: 10, 11, 15, 102, 104 and 108. In some embodiments the sarcoma-associated nucleic acid molecule comprises the nucleotide sequence set forth as

SEQ ID NO: 10. In other embodiments the sarcoma-associated antigen is encoded by a nucleic acid molecule comprising a nucleotide sequence set forth as SEQ ID NO: 133.

In yet other embodiments the sarcoma-associated antigen includes an amino acid sequence selected from the group consisting of SEQ ID NOs: 48, 50-53, 55-90, 111, 114, 116 and 120, or a fragment thereof that is at least eight amino acids in length. In still other embodiments the sarcoma-associated antigen includes an amino acid sequence selected from the group consisting of SEQ ID NOs: 50-52, 55-58, 60-90, 114, 116 and 120, or a fragment thereof that is at least eight amino acids in length. In some embodiments the sarcoma-associated antigen includes the amino acid sequence set forth as SEQ ID NO: 55, or a fragment thereof that is at least eight amino acids in length. In certain embodiments the sarcoma-associated antigen includes the amino acid sequence set forth as SEQ ID NO: 122. In still other embodiments the sarcoma-associated antigen includes the amino acid sequence set forth as SEQ ID NO: 124. In yet other embodiments the sarcoma-associated antigen includes the amino acid sequence set forth as SEQ ID NO: 126. In still other embodiments the sarcoma-associated antigen includes the amino acid sequence set forth as SEQ ID NOs: 128, 130 or 132. In still further embodiments the sarcoma-associated antigen includes the amino acid sequence set forth as SEQ ID NO: 134, or a fragment thereof that is at least eight amino acids in length.

In some embodiments, the agent that stimulates an immune response is a nucleic acid that encodes a sarcoma-associated antigen operably linked to a promoter for expressing the sarcoma-associated antigen; a polypeptide comprising the sarcoma-associated antigen; or a host cell that expresses the sarcoma-associated antigen, particularly a host cell that also expresses a MHC molecule. In some embodiments, the agent which stimulates an immune response is a peptide fragment of the sarcoma-associated antigen, or is a complex of a peptide fragment of the sarcoma-associated antigen and a MHC molecule. In other embodiments, the agent also includes an adjuvant or cytokine.

In another aspect of the invention, kits for diagnosing a disorder associated with the aberrant expression of a sarcoma-associated antigen or a nucleic acid molecule that encodes it are provided. The kits include one or more nucleic acid molecules that hybridize to the nucleic acid molecule that encodes the sarcoma-associated antigen comprising a nucleotide sequence selected from the group consisting of nucleotide sequences set forth as SEQ ID NOs: 3, 5-8, 10-45, 99, 102, 104 and 108 under high stringency conditions, and instructions for the use of the nucleic acid molecules in the diagnosis of a disorder associated with

aberrant expression of the sarcoma-associated antigen or the nucleic acid molecule that encodes it. In some embodiments the one or more nucleic acid molecules that hybridize to the nucleic acid molecule that encodes the sarcoma-associated antigen comprises a nucleotide sequence selected from the group consisting of nucleotide sequences set forth as SEQ ID NOs: 5-7, 10-13, 15-45, 102, 104 and 108. In some embodiments, the one or more nucleic acid molecules are detectably labeled. In some embodiments the nucleic acid molecule that encodes the sarcoma-associated antigen comprises the nucleotide sequence set forth as SEQ ID NO: 10, 11, 15, 102, 104 or 108. In other embodiments the nucleic acid molecule that encodes the sarcoma-associated antigen comprises the nucleotide sequence set forth as SEQ ID NO: 10. In other embodiments the nucleic acid molecule includes the nucleotide sequence set forth as SEQ ID NO: 121. In still other embodiments the nucleic acid molecule includes the nucleotide sequence set forth in SEQ ID NO: 123. In yet other embodiments the nucleic acid molecule includes the nucleotide sequence set forth in SEQ ID NO: 125. In still other embodiments the nucleic acid molecule includes the nucleotide sequence set forth in SEQ ID NOs: 127, 129 or 131.

In certain embodiments, the one or more nucleic acid molecules consist of a first primer and a second primer, wherein the first primer and the second primer are constructed and arranged to selectively amplify at least a portion of a nucleic acid molecule that encodes the sarcoma-associated antigen and comprises a nucleotide sequence selected from the group consisting of SEQ ID NOs: 10, 11, 15, 102, 104 and 108. In other embodiments, the nucleic acids in the kit are bound to a substrate.

In still another aspect of the invention, methods for identifying a cancer-associated antigen are provided. The methods include obtaining a biological sample from one or more subjects, determining the reactivity of the biological sample to one or more known cancer-associated antigens, using the reactive biological sample to screen an expression library to determine the presence of cancer-associated antigens reactive with the biological sample, and isolating a clone that encodes the cancer-associated antigen from the expression library. In certain embodiments the biological sample is serum. In some embodiments the expression library is derived from a tumor, preferably from a tumor cell line.

In still other embodiments, the methods also include determining the identity of the cancer-associated antigen encoded by the isolated clone, preferably by DNA sequencing.

The invention in a further aspect provides a composition including an agent that stimulates an immune response to a sarcoma-associated antigen. In some embodiments

sarcoma-associated antigens are those encoded by a nucleic acid molecule selected from the group consisting of an isolated nucleic acid molecule comprising a nucleotide sequence that is at least 90% identical to the nucleotide sequence selected from the group consisting of SEQ ID NOs: 3, 5-8, 10-45, 99, 102, 104 and 108, and nucleic acid molecules that differ from the nucleic acid molecules of (a) in codon sequence due to the degeneracy of the genetic code. In some embodiments the sarcoma-associated antigens are those encoded by a nucleic acid molecule selected from the group consisting of an isolated nucleic acid molecule comprising a nucleotide sequence that is at least 90% identical to the nucleotide sequence selected from the group consisting of SEQ ID NOs: 5-7, 10-13, 15-45, 102, 104 and 108, and nucleic acid molecules that differ from the nucleic acid molecules of (a) in codon sequence due to the degeneracy of the genetic code. In particular embodiments, the nucleic acid molecule includes a nucleotide sequence selected from the group consisting of SEQ ID NOs: 10, 11, 15, 102, 104 and 108. In some embodiments the nucleic acid molecule includes the nucleotide sequence set forth as SEQ ID NO: 10. In other embodiments the nucleic acid molecule includes the nucleotide sequence set forth as SEQ ID NO: 133.

In some embodiments, sarcoma-associated antigen comprises a polypeptide sequence selected from the group consisting of SEQ ID NOs: 48, 50-53, 55-90, 111, 114, 116 and 120 or a fragment thereof that is at least eight amino acids in length. In other embodiments, sarcoma-associated antigen comprises a polypeptide sequence selected from the group consisting of SEQ ID NOs: 50-52, 55-58, 60-90, 114, 116 and 120 or a fragment thereof that is at least eight amino acids in length. In some embodiments the sarcoma-associated antigen includes the amino acid sequence of SEQ ID NO: 55, 56, 60, 114, 116 or 120 or a fragment thereof that is at least eight amino acids in length. In other embodiments the sarcoma-associated antigen includes the amino acid sequence of SEQ ID NO: 55 or a fragment thereof that is at least eight amino acids in length. In certain embodiments the sarcoma-associated antigen includes the amino acid sequence set forth as SEQ ID NO: 122. In still other embodiments the sarcoma-associated antigen includes the amino acid sequence set forth as SEQ ID NO: 124. In yet other embodiments the sarcoma-associated antigen includes the amino acid sequence set forth as SEQ ID NO: 126. In still other embodiments the sarcoma-associated antigen includes the amino acid sequence set forth as SEQ ID NOs: 128, 130 or 132. In yet other embodiments the sarcoma-associated antigen includes the amino acid sequence set forth as SEQ ID NO: 134.

The agent, in some embodiments, is a nucleic acid that encodes a sarcoma-associated antigen operably linked to a promoter for expressing the sarcoma-associated antigen. In other embodiments, the agent is a polypeptide comprising the sarcoma-associated antigen. In still other embodiments, the agent is a host cell that expresses the sarcoma-associated antigen; preferably the host cell also expresses a MHC molecule. In yet other embodiments, the agent is a complex of a peptide derived from the sarcoma-associated antigen and a MHC molecule.

The composition also includes, in certain embodiments, an adjuvant or cytokine and/or one or more cytotoxic or chemotherapeutic agents. The compositions optionally includes a pharmaceutically acceptable carrier.

In another aspect of the invention, compositions are provided that include an agent that selectively binds to a sarcoma-associated antigen or a nucleic acid molecule that encodes it. The nucleic acid molecule includes a nucleotide sequence selected from the group consisting of: (a) an isolated nucleic acid molecule comprising a nucleotide sequence that is at least 90% identical to the nucleotide sequence selected from the group consisting of SEQ ID NOs: 3, 5-8, 10-13, 99, 102 and 104 and (b) nucleic acid molecules that differ from the nucleic acid molecules of (a) in codon sequence due to the degeneracy of the genetic code. In some embodiments the nucleic acid molecule includes a nucleotide sequence selected from the group consisting of: (a) an isolated nucleic acid molecule comprising a nucleotide sequence that is at least 90% identical to the nucleotide sequence selected from the group consisting of SEQ ID NOs: 5-7, 10-13, 102 and 104 and (b) nucleic acid molecules that differ from the nucleic acid molecules of (a) in codon sequence due to the degeneracy of the genetic code. In some embodiments the nucleic acid molecule includes a nucleotide sequence selected from the group consisting of SEQ ID NOs: 10, 11, 102 and 104; in other embodiments the nucleic acid molecule includes the nucleotide sequence set forth as SEQ ID NO: 10. In other embodiments the nucleic acid molecule includes the nucleotide sequence set forth as SEQ ID NO: 121. In still other embodiments the nucleic acid molecule includes the nucleotide sequence set forth in SEQ ID NO: 123. In yet other embodiments the nucleic acid molecule includes the nucleotide sequence set forth in SEQ ID NO: 125. In still other embodiments the nucleic acid molecule includes the nucleotide sequence set forth in SEQ ID NOs: 127, 129 or 131. In other embodiments, the sarcoma-associated antigen includes the amino acid sequence set forth as SEQ ID NO: 55, 56, 114 or 116 or a fragment thereof that is at least eight amino acids in length. In some embodiments the sarcoma-associated antigen includes the amino acid sequence set forth as SEQ ID NO: 55. In certain embodiments the

sarcoma-associated antigen includes the amino acid sequence set forth as SEQ ID NO: 122. In still other embodiments the sarcoma-associated antigen includes the amino acid sequence set forth as SEQ ID NO: 124. In yet other embodiments the sarcoma-associated antigen includes the amino acid sequence set forth as SEQ ID NO: 126. In still other embodiments the sarcoma-associated antigen includes the amino acid sequence set forth as SEQ ID NOs: 128, 130 or 132. The agents in this aspect of the invention include nucleic acids and polypeptides, preferably antibodies or antigen-binding fragments thereof. Preferred antibodies include monoclonal antibodies (particularly chimeric, human, or humanized antibodies), and single chain antibodies; preferred antibody fragments include $F(ab')_2$, Fab, Fd, or Fv fragments.

In certain embodiments, the antibody or antigen-binding fragment is conjugated to cytotoxic or chemotherapeutic agent. In other embodiments, the composition includes one or more cytotoxic or chemotherapeutic agent. In still other embodiments, the composition includes a pharmaceutically acceptable carrier.

The use of the nucleotide and amino acid sequence as set forth as SEQ ID NOs: 133 and 134, respectively, in any of the compositions and methods described herein are also provided.

The invention also involves the use of the genes, gene products, fragments thereof, agents which bind thereto, and other compositions and molecules described herein in the preparation of medicaments. A particular medicament is for treating cancer.

These and other aspects of the invention will be described in further detail in connection with the detailed description of the invention.

## Brief Description of the Drawing

Figure 1 provides the mRNA expression patterns of serologically defined sarcoma antigens. Figure 1A shows the results of the RT-PCR analysis of NY-SAR-12, -35, and –41 in a panel of 17 normal tissues (Lanes 1, brain; 2, kidney; 3, liver; 4, pancreas; 5, placenta; 6, testis; 7, fetal brain; 8, small intestine; 9, heart; 10, prostate; 11, adrenal gland; 12, spleen; 13, colon; 14, stomach; 15, lung; 16, bladder; and 17, ovary). Figure 1B provides the results of the quantitative real-time RT-PCR analysis of NY-SAR-35 in various normal tissues. Figure 1C shows the results of the RT-PCR analysis of NY-SAR-35 expression in sarcoma cell lines and sarcoma tissue (Lane 1, fibrosarcoma; 2, rhabdomyosarcoma; 3, leiomyosarcoma; and 4, normal testis). Figure 1D provides the results of the Northern blot analysis of NY-SAR-35 in

various normal tissues (Lane 1, spleen; 2, thymus; 3, prostate; 4, testis; 5, ovary; 6, small intestine; 7, colon mucosa; and 8, peripheral blood leukocytes).

Figure 2 provides the nucleotide and predicted amino acid sequence of NY-SAR-35 from each of the four ATG codons. The underlined letters indicate the signal peptide and the italicized letters indicate the transmembrane domain. The letters shown in gray represent the trefoil domain, while the letters that are underlined and italicized represent the other hydrophilic turn.

Figure 3 provides the results of Western blot assay of recombinant NY-SAR-35 proteins in *E. coli*. Three colonies of each domain cloned plasmid were picked and cultured by IPTG induction. After a four hour induction, total proteins from each of the colonies were separated by SDS-gel electrophoresis. The protein gel was immunoblotted on a membrane with a His-epitope monoclonal antibody. Lanes 1, 2, and 3-whole protein (from the first ATG codon); Lanes 4, 5 and 6 – MH7 protein; Lanes 7, 8, and 9-extracellular protein and Lanes 10 and 11-*E. coli* lysate as negative control.

Figure 4 provides the real-time RT-PCR analysis of NY-SAR-35 mRNA in various normal tissues and non-small cell lung cancer specimens. NY-SAR-35 was expressed in normal testis (83.2 ag) at a level that was >1,000 times the level detected in all other normal tissues. In 2 of 9 cases of non-small cell lung cancer examined, the level of NY-SAR-35 expression was equivalent to 0.15 (12.5 ag) and 0.13 (10.8 ag) times the level detected in normal testis, or approximately 100 times the level detected in normal tissues.

## Detailed Description of the Invention

The screening of cDNA expression libraries derived from human tumors with autologous antibody (SEREX) has proven to be a powerful method for defining the structure of tumor antigens recognized by the humoral immune system, and has led to the identification of new targets for cancer immunotherapy. The current study examined the humoral immune response of sarcoma patients to CT antigens. Sera from patients which showed a humoral immune response to CT antigens were subsequently used to screen cDNA libraries derived from CT-rich sarcoma cell lines, leading to the identification of antigens not before associated with cancer along with several novel antigens associated with a sarcoma-related immune response, including a novel CT antigen, NY-SAR-35.

Sarcoma-associated antigens were identified with an optimized SEREX analysis method. Cell lines that were rich in CT antigen expression were chosen as the source of

cDNA. Additionally, sera was obtained from a group of patients that were actively mounting a humoral immune response to a panel of known CT antigens. This optimized SEREX analysis led to the identification of 113 antigens reactive with serum IgG of sarcoma patients. The antigens identified were further evaluated for cancer-restricted expression and the frequency of eliciting antibody responses in normal individuals as well as cancer patients.

In the first round of immunoscreenings, twenty-four of 72 antigens (33%) were found to have a serological profile that was not restricted to cancer patients, as evidenced by their reactivity with normal sera, while 48 antigens had a cancer-related serological profile, reacting only with sera from cancer patients. Notable antigens belonging to this latter category include the CT antigens, NY-SAR-36/SSX-1, NY-SAR-43/SSX-4 and NY-SAR-35. Although the antibody response in these studies to NY-SAR-4/FH was most frequent, occurring in 5/39 (13%) sarcoma patients, no individual antigen was serodominant. NY-SAR-4 is equivalent to fumarate hydratase (FH), an enzyme of the tricarboxylic acid cycle. This serological response to NY-SAR-4/FH may be of interest given the recent finding that germ line mutations in the FH gene are associated with a predisposition to uterine and cutaneous leiomyomata, and also renal cell carcinoma (Tomlinson IP, et al. Germline mutations in FH predispose to dominantly inherited uterine fibroids, skin leiomyomata and papillary renal cell cancer. Nat Genet 2002 Apr;30(4):406-10).

In addition, 6 tissue-restricted antigens, LAGE-1/NY-SAR-17, SSX1/NY-SAR-36, SSX4/NY-SAR-43, NESG1/NY-SAR-12, NY-SAR-35, and NY-SAR-41 were identified. Two of these antigens, NY-SAR-35, and NY-SAR-41 are novel gene products, and a third, NESG1/NY-SAR-12 (Li Z, Yao K, Cao Y. Molecular cloning of a novel tissue-specific gene from human nasopharyngeal epithelium. Gene 1999 Sep 3;237(1):235-40), has not been previously studied in relation to cancer. NY-SAR-35 further represents a newly defined CT antigen expressed exclusively in normal testis, melanoma, sarcoma, lung cancer and breast cancer.

The second round of immunoscreenings performed led to the identification of 41 additional SEREX-defined sarcoma antigens, 11 of which are novel gene products (NY-SAR-77, -79, -80, -84, -88, -92, -95, -97, -104, 105 and -113). Within this group of 41 sarcoma antigens are three known testis-restricted antigens (NY-SAR-78/TSP-NY, NY-SAR-89/SSX2 and NY-SAR-99/SSX3), two differentially expressed antigens that are novel gene products (NY-SAR-92 and NY-SAR-97) and a tissue-restricted antigen that has not been previously studied in relation to cancer (NY-SAR-96/MCSP).

Table 1, below, provides a list of the sarcoma-associated antigens and their corresponding sequence identification numbers. The antigens listed include those that were found to be uncharacterized gene products as well as those sarcoma-associated antigens that exhibited cancer-restricted expression and were not found in the SEREX Database.

**Table 1: Sarcoma-Associated Antigens (Uncharacterized Gene Products and Cancer-Related Antigens not Found in the SEREX Database)**

| NY-SAR-Antigen | Sequence Identification Number (nucleotide and amino acid sequence, respectively) | NY-SAR-Antigen | Sequence Identification Number (nucleotide and amino acid sequence, respectively |
|---|---|---|---|
| 3 | SEQ ID NOs: 1 and 46 | 40 | SEQ ID NOs: 30 and 75 |
| 10 | SEQ ID NOs: 2 and 47 | 42 | SEQ ID NOs: 31 and 76 |
| 16 | SEQ ID NOs: 3 and 48 | 43 | SEQ ID NOs: 32 and 77 |
| 22 | SEQ ID NOs: 4 and 49 | 46 | SEQ ID NOs: 33 and 78 |
| 23 | SEQ ID NOs: 5 and 50 | 49 | SEQ ID NOs: 34 and 79 |
| 24 | SEQ ID NOs: 6 and 51 | 50 | SEQ ID NOs: 35 and 80 |
| 27 | SEQ ID NOs: 7 and 52 | 51 | SEQ ID NOs: 36 and 81 |
| 28 | SEQ ID NOs: 8 and 53 | 52 | SEQ ID NOs: 37 and 82 |
| 29 | SEQ ID NOs: 9 and 54 | 56 | SEQ ID NOs: 38 and 83 |
| 35 | SEQ ID NOs: 10 and 55 | 57 | SEQ ID NOs: 39 and 84 |
| 41 | SEQ ID NOs: 11 and 56 | 59 | SEQ ID NOs: 40 and 85 |
| 48 | SEQ ID NOs: 12 and 57 | 60 | SEQ ID NOs: 41 and 86 |
| 62 | SEQ ID NOs: 13 and 58 | 63 | SEQ ID NOs: 42 and 87 |
| 71 | SEQ ID NOs: 14 and 59 | 67 | SEQ ID NOs: 43 and 88 |
| 12 | SEQ ID NOs: 15 and 60 | 69 | SEQ ID NOs: 44 and 89 |
| 4 | SEQ ID NOs: 16 and 61 | 70 | SEQ ID NOs: 45 and 90 |
| 5 | SEQ ID NOs: 17 and 62 | 77 | SEQ ID NOs: 97 and 109 |
| 8 | SEQ ID NOs: 18 and 63 | 79 | SEQ ID NOs: 98 and 110 |
| 9 | SEQ ID NOs: 19 and 64 | 80 | SEQ ID NOs: 99 and 111 |
| 20 | SEQ ID NOs: 20 and 65 | 84 | SEQ ID NOs: 100 and 112 |
| 21 | SEQ ID NOs: 21 and 66 | 88 | SEQ ID NOs: 101 and 113 |
| 25 | SEQ ID NOs: 22 and 67 | 92 | SEQ ID NOs: 102 and 114 |
| 26 | SEQ ID NOs: 23 and 68 | 95 | SEQ ID NOs: 103 and 115 |
| 30 | SEQ ID NOs: 24 and 69 | 97 | SEQ ID NOs: 104 and 116 |
| 34 | SEQ ID NOs: 25 and 70 | 104 | SEQ ID NOs: 105 and 117 |
| 36 | SEQ ID NOs: 26 and 71 | 105 | SEQ ID NOs: 106 and 118 |
| 37 | SEQ ID NOs: 27 and 72 | 113 | SEQ ID NOs: 107 and 119 |
| 38 | SEQ ID NOs: 28 and 73 | 96 | SEQ ID NOs: 108 and 120 |
| 39 | SEQ ID NOs: 29 and 74 | | |

The invention relates, in part, to the sarcoma-associated antigens defined herein and the nucleic acid molecules that encode them. The invention further relates to the use of the nucleic acid molecules, polypeptides and fragments thereof associated with sarcoma in methods and compositions for the diagnosis and treatment of diseases, such as cancer.

As used herein, the term "sarcoma-associated antigens" means polypeptides that elicit specific immune responses to the polypeptide when expressed by a tumor cell and thus, include sarcoma-associated polypeptides (including proteins) and fragments of sarcoma-associated polypeptides, that are recognized by the immune system (e.g., by antibodies and/or T lymphocytes). In part, the invention relates to sarcoma-associated antigens as well as the nucleic acid molecules that encode the sarcoma-associated antigens. As used herein, the "nucleic acid molecules that encode" means the nucleic acid molecules that code for the immunogenic sarcoma-associated polypeptides or immunogenic fragments thereof. These nucleic acid molecules may be DNA or may be RNA (e.g. mRNA). The sarcoma-associated nucleic acid molecules of the invention also encompass variants of the nucleic acid molecules described herein. These variants may be splice variants or allelic variants of certain sequences provided. Variants of the nucleic acid molecules of the invention are intended to include homologs and alleles which are described further below. Further, as used herein, the term "sarcoma-associated molecules" includes sarcoma-associated antigens (polypeptides and fragments thereof) as well as sarcoma-associated nucleic acids. In all embodiments, human sarcoma-associated antigens and the encoding nucleic acid molecules thereof, are preferred.

In one aspect, the invention provides isolated nucleic acid molecules that encode the sarcoma-associated antigens defined herein. The isolated nucleic acid molecules of this aspect of the invention comprise: (a) nucleotide sequences selected from the group consisting of nucleotide sequences set forth as SEQ ID NOs: 1-14 and 97-107 (b) isolated nucleic acid molecules which hybridize under highly stringent conditions to the nucleic acid molecules of (a) and which code for a sarcoma-associated antigen, (c) nucleic acid molecules that differ from (a) or (b) due to the degeneracy of the genetic code, and (d) complements of (a), (b) or (c).

As used herein the term "isolated nucleic acid molecule" means: (i) amplified *in vitro* by, for example, polymerase chain reaction (PCR); (ii) recombinantly produced by cloning; (iii) purified, as by cleavage and gel separation; or (iv) synthesized by, for example, chemical synthesis. An isolated nucleic acid is one which is readily manipulable by recombinant DNA

techniques well known in the art. Thus, a nucleotide sequence contained in a vector in which 5' and 3' restriction sites are known or for which polymerase chain reaction (PCR) primer sequences have been disclosed is considered isolated but a nucleic acid sequence existing in its native state in its natural host is not. An isolated nucleic acid may be substantially purified, but need not be. For example, a nucleic acid that is isolated within a cloning or expression vector is not pure in that it may comprise only a tiny percentage of the material in the cell in which it resides. Such a nucleic acid is isolated, however, as the term is used herein because it is readily manipulable by standard techniques known to those of ordinary skill in the art.

The sarcoma-associated nucleic acid molecules of the invention also intended to encompass homologs and alleles which can be identified by conventional techniques. Identification of human and other organism homologs of sarcoma-associated polypeptides will be familiar to those of skill in the art. In general, nucleic acid hybridization is a suitable method for identification of homologous sequences of another species (e.g., human, cow, sheep), which correspond to a known sequence. Standard nucleic acid hybridization procedures can be used to identify related nucleic acid sequences of selected percent identity. For example, one can construct a library of cDNAs reverse transcribed from the mRNA of a selected tissue and use the nucleic acids that encode sarcoma-associated antigens identified herein to screen the library for related nucleotide sequences. The screening preferably is performed using high-stringency conditions to identify those sequences that are closely related by sequence identity. Nucleic acids so identified can be translated into polypeptides and the polypeptides can be tested for activity.

The term "high stringency" as used herein refers to parameters with which the art is familiar. Nucleic acid hybridization parameters may be found in references that compile such methods, e.g. Molecular Cloning: A Laboratory Manual, J. Sambrook, et al., eds., Second Edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 1989, or Current Protocols in Molecular Biology, F.M. Ausubel, et al., eds., John Wiley & Sons, Inc., New York. More specifically, high-stringency conditions, as used herein, refers, for example, to hybridization at 65°C in hybridization buffer (3.5X SSC, 0.02% Ficoll, 0.02% polyvinyl pyrrolidone, 0.02% Bovine Serum Albumin, 2.5mM NaH2PO4(pH7), 0.5% SDS, 2mM EDTA). SSC is 0.15M sodium chloride/0.15M sodium citrate, pH7; SDS is sodium dodecyl sulphate; and EDTA is ethylenediaminetetracetic acid. After hybridization, the

membrane upon which the DNA is transferred is washed, for example, in 2X SSC at room temperature and then at 0.1 - 0.5X SSC/0.1X SDS at temperatures up to 68°C.

There are other conditions, reagents, and so forth that can be used, which result in a similar degree of stringency. The skilled artisan will be familiar with such conditions, and thus they are not given here. It will be understood, however, that the skilled artisan will be able to manipulate the conditions in a manner to permit the clear identification of homologs and alleles of the sarcoma-associated nucleic acids of the invention (e.g., by using lower stringency conditions). The skilled artisan also is familiar with the methodology for screening cells and libraries for expression of such molecules, which then are routinely isolated, followed by isolation of the pertinent nucleic acid molecule and sequencing.

In general, homologs and alleles typically will share at least 90% nucleotide identity and/or at least 95% amino acid identity to the sequences of sarcoma-associated nucleic acids and polypeptides, respectively, in some instances will share at least 95% nucleotide identity and/or at least 97% amino acid identity, in other instances will share at least 97% nucleotide identity and/or at least 98% amino acid identity, in other instances will share at least 99% nucleotide identity and/or at least 99% amino acid identity, and in other instances will share at least 99.5% nucleotide identity and/or at least 99.5% amino acid identity. The homology can be calculated using various, publicly available software tools developed by NCBI (Bethesda, Maryland) that can be obtained through the internet. Exemplary tools include the BLAST system available from the website of the National Center for Biotechnology Information (NCBI) at the National Institutes of Health. Pairwise and ClustalW alignments (BLOSUM30 matrix setting) as well as Kyte-Doolittle hydropathic analysis can be obtained using the MacVector sequence analysis software (Oxford Molecular Group). Watson-Crick complements of the foregoing nucleic acids also are embraced by the invention.

In another aspect of the invention, unique fragments are provided which include unique fragments of the nucleotide sequences of the invention and complements thereof. The invention, in a preferred embodiment, provides unique fragments of SEQ ID NO: 10, 11, 15, 102, 104 or 108 and complements thereof. In another preferred embodiment, provides unique fragments of SEQ ID NO: 10 and complements thereof. In other embodiments the unique fragment includes the nucleotide sequence set forth as SEQ ID NO: 121. In still other embodiments the unique fragments includes the sequence set forth as SEQ ID NO: 123, 125, 127, 129 or 131. A unique fragment is one that is a 'signature' for the larger nucleic acid. It, for example, is long enough to assure that its precise sequence is not found in molecules

outside of the nucleic acid molecules that encode the sarcoma-associated antigens defined above. Those of ordinary skill in the art may apply no more than routine procedures to determine if a fragment is unique within the human genome. In some instances the unique fragment is at least about 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 35, 40, 45, 50, 75, or 100 amino acids in length.

Unique fragments can be used as probes in Southern blot assays to identify such nucleic acid molecules, or can be used as probes in amplification assays such as those employing the polymerase chain reaction (PCR), including, but not limited to RT-PCR and RT-real-time PCR. As known to those skilled in the art, large probes such as 200 nucleotides or more are preferred for certain uses such as Southern blots, while smaller fragments will be preferred for uses such as PCR. Unique fragments also can be used to produce fusion proteins for generating antibodies or determining binding of the polypeptide fragments, or for generating immunoassay components. Likewise, unique fragments can be employed to produce nonfused fragments of the sarcoma-associated polypeptides useful, for example, in the preparation of antibodies and in immunoassays.

In screening for sarcoma-associated antigen genes, a Southern blot may be performed using the foregoing conditions, together with a detectably labeled probe (e.g. radioactive or chemiluminescent probes). After washing the membrane to which the DNA is finally transferred, the membrane can be placed against X-ray film or analyzed using a phosphorimager device to detect the radioactive or chemiluminescent signal. In screening for the expression of sarcoma-associated antigen nucleic acids, Northern blot hybridizations using the foregoing conditions can be performed on samples taken from cancer patients or subjects suspected of having a condition characterized by abnormal cell proliferation or neoplasia. Amplification protocols such as polymerase chain reaction using primers that hybridize to the sequences presented also can be used for detection of the sarcoma-associated antigen genes or expression thereof.

Identification of related sequences can also be achieved using polymerase chain reaction (PCR) and other amplification techniques suitable for cloning related nucleic acid sequences. Preferably, PCR primers are selected to amplify portions of a nucleic acid sequence believed to be conserved (e.g., a catalytic domain, a DNA-binding domain, etc.). Again, nucleic acids are preferably amplified from a tissue-specific library (e.g., testis). One also can use expression cloning utilizing the antisera described herein to identify nucleic acids that encode related antigenic proteins in humans or other species using the SEREX

procedure to screen the appropriate expression libraries. (See: Sahin et al. Proc. Natl. Acad. Sci. USA 92:11810-11813, 1995).

The invention also includes degenerate nucleic acids that include alternative codons to those present in the native materials. For example, serine residues are encoded by the codons TCA, AGT, TCC, TCG, TCT and AGC. Each of the six codons is equivalent for the purposes of encoding a serine residue. Thus, it will be apparent to one of ordinary skill in the art that any of the serine-encoding nucleotide triplets may be employed to direct the protein synthesis apparatus, in vitro or in vivo, to incorporate a serine residue into an elongating sarcoma-associated polypeptide. Similarly, nucleotide sequence triplets which encode other amino acid residues include, but are not limited to: CCA, CCC, CCG, and CCT (proline codons); CGA, CGC, CGG, CGT, AGA, and AGG (arginine codons); ACA, ACC, ACG, and ACT (threonine codons); AAC and AAT (asparagine codons); and ATA, ATC, and ATT (isoleucine codons). Other amino acid residues may be encoded similarly by multiple nucleotide sequences. Thus, the invention embraces degenerate nucleic acids that differ from the biologically isolated nucleic acids in codon sequence due to the degeneracy of the genetic code.

The invention also provides modified nucleic acid molecules, which include additions, substitutions and deletions of one or more nucleotides (preferably 1-20 nucleotides). In preferred embodiments, these modified nucleic acid molecules and/or the polypeptides they encode retain at least one activity or function of the unmodified nucleic acid molecule and/or the polypeptides, such as antigenicity, receptor binding, etc. In certain embodiments, the modified nucleic acid molecules encode modified polypeptides, preferably polypeptides having conservative amino acid substitutions as are described elsewhere herein. The modified nucleic acid molecules are structurally related to the unmodified nucleic acid molecules and in preferred embodiments are sufficiently structurally related to the unmodified nucleic acid molecules so that the modified and unmodified nucleic acid molecules hybridize under stringent conditions known to one of skill in the art.

For example, modified nucleic acid molecules that encode polypeptides having single amino acid changes can be prepared. Each of these nucleic acid molecules can have one, two or three nucleotide substitutions exclusive of nucleotide changes corresponding to the degeneracy of the genetic code as described herein. Likewise, modified nucleic acid molecules that encode polypeptides having two amino acid changes can be prepared which have, e.g., 2-6 nucleotide changes. Numerous modified nucleic acid molecules like these will

be readily envisioned by one of skill in the art, including for example, substitutions of nucleotides in codons encoding amino acids 2 and 3, 2 and 4, 2 and 5, 2 and 6, and so on. In the foregoing example, each combination of two amino acids is included in the set of modified nucleic acid molecules, as well as all nucleotide substitutions which code for the amino acid substitutions. Additional nucleic acid molecules that encode polypeptides having additional substitutions (i.e., 3 or more), additions or deletions (e.g., by introduction of a stop codon or a splice site(s)) also can be prepared and are embraced by the invention as readily envisioned by one of ordinary skill in the art. Any of the foregoing nucleic acids or polypeptides can be tested by routine experimentation for retention of activity or structural relation to the nucleic acids and/or polypeptides disclosed herein. As used herein the terms: "deletion", "addition", and "substitution" mean deletion, addition, and substitution changes to about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, or more nucleic acids of a sequence of the invention.

According to yet another aspect of the invention, an expression vector comprising any of the isolated nucleic acid molecules of the invention, preferably operably linked to a promoter is provided. In a related aspect, host cells transformed or transfected with such expression vectors also are provided. As used herein, a "vector" may be any of a number of nucleic acid molecules into which a desired sequence may be inserted by restriction and ligation for transport between different genetic environments or for expression in a host cell. Vectors are typically composed of DNA although RNA vectors are also available. Vectors include, but are not limited to, plasmids, phagemids, and virus genomes. A cloning vector is one which is able to replicate in a host cell, and which is further characterized by one or more endonuclease restriction sites at which the vector may be cut in a determinable fashion and into which a desired DNA sequence may be ligated such that the new recombinant vector retains its ability to replicate in the host cell. In the case of plasmids, replication of the desired sequence may occur many times as the plasmid increases in copy number within the host bacterium or just a single time per host before the host reproduces by mitosis. In the case of phage, replication may occur actively during a lytic phase or passively during a lysogenic phase. An expression vector is one into which a desired DNA sequence may be inserted by restriction and ligation such that it is operably joined to regulatory sequences and may be expressed as an RNA transcript. Vectors may further contain one or more marker sequences suitable for use in the identification of cells which have or have not been transformed or transfected with the vector. Markers include, for example, genes encoding

proteins which increase or decrease either resistance or sensitivity to antibiotics or other compounds, genes which encode enzymes whose activities are detectable by standard assays known in the art, e.g., -galactosidase or alkaline phosphatase, and genes which visibly affect the phenotype of transformed or transfected cells, hosts, colonies or plaques, e.g., green fluorescent protein. Preferred vectors are those capable of autonomous replication and expression of the structural gene products present in the DNA segments to which they are operably joined.

As used herein, a coding sequence and regulatory sequences are said to be "operably joined" when they are covalently linked in such a way as to place the expression or transcription of the coding sequence under the influence or control of the regulatory sequences. As used herein, "operably joined" and "operably linked" are used interchangeably and should be construed to have the same meaning. If it is desired that the coding sequences be translated into a functional protein, two DNA sequences are said to be operably joined if induction of a promoter in the 5' regulatory sequences results in the transcription of the coding sequence and if the nature of the linkage between the two DNA sequences does not (1) result in the introduction of a frame-shift mutation, (2) interfere with the ability of the promoter region to direct the transcription of the coding sequences, or (3) interfere with the ability of the corresponding RNA transcript to be translated into a protein. Thus, a promoter region is operably joined to a coding sequence if the promoter region is capable of effecting transcription of that DNA sequence such that the resulting transcript can be translated into the desired protein or polypeptide.

The precise nature of the regulatory sequences needed for gene expression may vary between species or cell types, but shall in general include, as necessary, 5' non-transcribed and 5' non-translated sequences involved with the initiation of transcription and translation respectively, such as a TATA box, capping sequence, CAAT sequence, and the like. Often, such 5' non-transcribed regulatory sequences will include a promoter region which includes a promoter sequence for transcriptional control of the operably joined gene. Regulatory sequences may also include enhancer sequences or upstream activator sequences as desired. The vectors of the invention may optionally include 5' leader or signal sequences. The choice and design of an appropriate vector is within the ability and discretion of one of ordinary skill in the art.

It will also be recognized that the invention embraces the use of the sarcoma-associated nucleic acid molecules and genomic sequences in expression vectors, as well as to

transfect host cells and cell lines, be these prokaryotic, e.g., *E. coli*, or eukaryotic, e.g., CHO cells, COS cells, yeast expression systems, and recombinant baculovirus expression in insect cells. Especially useful are mammalian cells such as human, mouse, hamster, pig, goat, primate, etc. They may be of a wide variety of tissue types, including mast cells, fibroblasts, oocytes, and lymphocytes, and may be primary cells and cell lines. Specific examples include dendritic cells, peripheral blood leukocytes, bone marrow stem cells and embryonic stem cells. The expression vectors require that the pertinent sequence, i.e., those nucleic acids described supra, be operably linked to a promoter.

The invention, in one aspect, also permits the construction of sarcoma-associated antigen gene "knock-outs" and "knock-ins" in cells and in animals, providing materials for studying certain aspects of cancer and immune system responses to cancer.

Expression vectors containing all the necessary elements for expression are commercially available and known to those skilled in the art. See, e.g., Sambrook et al., Molecular Cloning: A Laboratory Manual, Second Edition, Cold Spring Harbor Laboratory Press, 1989. Cells are genetically engineered by the introduction into the cells of heterologous DNA or RNA encoding a sarcoma-associated antigen, a mutant sarcoma-associated antigen, fragments, or variants thereof. The heterologous DNA or RNA is placed under operable control of transcriptional elements to permit the expression of the heterologous DNA in the host cell.

Preferred systems for mRNA expression in mammalian cells are those such as pcDNA1.1 and pCDM8 (Invitrogen) that contain a selectable marker (which facilitates the selection of stably transfected cell lines) and contain the human cytomegalovirus (CMV) enhancer-promoter sequences. Additionally, suitable for expression in primate or canine cell lines is the pCEP4 vector (Invitrogen), which contains an Epstein Barr virus (EBV) origin of replication, facilitating the maintenance of plasmid as a multicopy extrachromosomal element. Another expression vector is the pEF-BOS plasmid containing the promoter of polypeptide Elongation Factor 1, which stimulates efficiently transcription in vitro. The plasmid is described by Mizushima and Nagata (Nuc. Acids Res. 18:5322, 1990), and its use in transfection experiments is disclosed by, for example, Demoulin (Mol. Cell. Biol. 16:4710-4716, 1996). Still another preferred expression vector is an adenovirus, described by Stratford-Perricaudet, which is defective for E1 and E3 proteins (J. Clin. Invest. 90:626-630, 1992). The use of the adenovirus as an Adeno.P1A recombinant is described by Warnier et

al., in intradermal injection in mice for immunization against P1A (Int. J. Cancer, 67:303-310, 1996).

The invention also embraces kits termed expression kits, which allow the artisan to prepare a desired expression vector or vectors. Such expression kits include at least separate portions of each of the previously discussed coding sequences. Other components may be added, as desired, as long as the previously mentioned sequences, which are required, are included.

The invention also includes kits for amplification of a sarcoma-associated antigen nucleic acid, including at least one pair of amplification primers which hybridize to a sarcoma-associated nucleic acid. The primers preferably are about 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31 or 32 nucleotides in length and are non-overlapping to prevent formation of "primer-dimers". One of the primers will hybridize to one strand of the sarcoma-associated nucleic acid and the second primer will hybridize to the complementary strand of the sarcoma-associated nucleic acid, in an arrangement which permits amplification of the sarcoma-associated nucleic acid. Selection of appropriate primer pairs is standard in the art. For example, the selection can be made with assistance of a computer program designed for such a purpose, optionally followed by testing the primers for amplification specificity and efficiency.

The invention, in another aspect provides isolated polypeptides (including whole proteins and partial proteins) encoded by the foregoing sarcoma-associated nucleic acids. Examples of the amino acid sequences encoded by the foregoing sarcoma-associated nucleic acids are set forth as SEQ ID NOs: 46-90 and 109-120. The amino acids of the invention are also intended to encompass amino acid sequences that result from the translation of the nucleic acid sequences provided herein in a different reading frame. In one preferred embodiment of the invention a polypeptide is provided which comprises the polypeptide sequence set forth as SEQ ID NO: 55, 56, 60, 114, 116 or 120. In another preferred embodiment a polypeptide is provided which comprises the polypeptide sequence set forth as SEQ ID NO: 122. In still another preferred embodiment a polypeptide is provided which comprises the polypeptide sequence set forth as SEQ ID NO: 124. In still other embodiments polypeptides are provided which comprise the polypeptide sequence set forth as SEQ ID NO: 126, 128, 130 or 132. Such polypeptides are useful, for example, alone or as fusion proteins to generate antibodies, and as components of an immunoassay or diagnostic assay. Immunogenic sarcoma-associated polypeptides can be isolated from biological samples

including tissue or cell homogenates, and can also be expressed recombinantly in a variety of prokaryotic and eukaryotic expression systems by constructing an expression vector appropriate to the expression system, introducing the expression vector into the expression system, and isolating the recombinantly expressed protein. Fragments of the immunogenic sarcoma-associated polypeptides (including immunogenic peptides) also can be synthesized chemically using well-established methods of peptide synthesis. Thus, fragments of the disclosed polypeptides are useful for eliciting an immune response. In one embodiment fragments of a polypeptide which comprises SEQ ID NO: 55, 56, 60, 114, 116 or 120 that are at least eight amino acids in length and exhibit immunogenicity are provided. In one embodiment fragments of a polypeptide which comprises SEQ ID NO: 55 that are at least eight amino acids in length and exhibit immunogenicity are provided. In another embodiment a polypeptide is provided which comprises the polypeptide sequence set forth as SEQ ID NO: 122. In still another preferred embodiments a polypeptide is provided which comprises the polypeptide sequence set forth as SEQ ID NO: 124. In still other embodiments polypeptides are provided which comprise the polypeptide sequence set forth as SEQ ID NO: 126, 128, 130 or 132.

Fragments of a polypeptide preferably are those fragments that retain a distinct functional capability of the polypeptide. Functional capabilities that can be retained in a fragment of a polypeptide include interaction with antibodies or MHC molecules (e.g. immunogenic fragments), interaction with other polypeptides or fragments thereof, selective binding of nucleic acids or proteins, and enzymatic activity. One important activity is the ability to provoke in a subject an immune response. As will be recognized by those skilled in the art, the size of the fragment that can be used for inducing an immune response will depend upon factors such as whether the epitope recognized by an antibody is a linear epitope or a conformational epitope or the particular MHC molecule that binds to and presents the fragment (e.g. HLA class I or II). Thus, some immunogenic fragments of sarcoma-associated polypeptides will consist of longer segments while others will consist of shorter segments, (e.g. about 5, 6, 7, 8, 9, 10, 11 or 12 or more amino acids long, including each integer up to the full length of the sarcoma-associated polypeptide). Those skilled in the art are well versed in methods for selecting immunogenic fragments of polypeptides.

The invention embraces variants of the sarcoma-associated polypeptides described above. As used herein, a "variant" of a sarcoma-associated antigen polypeptide is a polypeptide which contains one or more modifications to the primary amino acid sequence of

a sarcoma-associated polypeptide. Modifications which create a sarcoma-associated antigen variant can be made to a sarcoma-associated polypeptide 1) to reduce or eliminate an activity of a sarcoma-associated polypeptide; 2) to enhance a property of a sarcoma-associated polypeptide, such as protein stability in an expression system or the stability of protein-protein binding; 3) to provide a novel activity or property to a sarcoma-associated polypeptide, such as addition of an antigenic epitope or addition of a detectable moiety; or 4) to provide equivalent or better binding to a MHC molecule.

Modifications to a sarcoma-associated polypeptide are typically made to the nucleic acid which encodes the sarcoma-associated polypeptide, and can include deletions, point mutations, truncations, amino acid substitutions and additions of amino acids or non-amino acid moieties. Alternatively, modifications can be made directly to the polypeptide, such as by cleavage, addition of a linker molecule, addition of a detectable moiety, such as biotin, addition of a fatty acid, and the like. Modifications also embrace fusion proteins comprising all or part of the sarcoma-associated antigen amino acid sequence. One of skill in the art will be familiar with methods for predicting the effect on protein conformation of a change in protein sequence, and can thus "design" a variant sarcoma-associated polypeptide according to known methods. One example of such a method is described by Dahiyat and Mayo in *Science* 278:82-87, 1997, whereby proteins can be designed *de novo*. The method can be applied to a known protein to vary a only a portion of the polypeptide sequence. By applying the computational methods of Dahiyat and Mayo, specific variants of a sarcoma-associated polypeptide can be proposed and tested to determine whether the variant retains a desired conformation.

In general, variants include sarcoma-associated polypeptides which are modified specifically to alter a feature of the polypeptide unrelated to its desired physiological activity. For example, cysteine residues can be substituted or deleted to prevent unwanted disulfide linkages. Similarly, certain amino acids can be changed to enhance expression of a sarcoma-associated polypeptide by eliminating proteolysis by proteases in an expression system (e.g., dibasic amino acid residues in yeast expression systems in which KEX2 protease activity is present).

Mutations of a nucleic acid which encode a sarcoma-associated polypeptide preferably preserve the amino acid reading frame of the coding sequence, and preferably do not create regions in the nucleic acid which are likely to hybridize to form secondary

structures, such a hairpins or loops, which can be deleterious to expression of the variant polypeptide.

Mutations can be made by selecting an amino acid substitution, or by random mutagenesis of a selected site in a nucleic acid which encodes the polypeptide. Variant polypeptides are then expressed and tested for one or more activities to determine which mutation provides a variant polypeptide with the desired properties. Further mutations can be made to variants (or to non-variant sarcoma-associated polypeptides) which are silent as to the amino acid sequence of the polypeptide, but which provide preferred codons for translation in a particular host. The preferred codons for translation of a nucleic acid in, e.g., *E. coli*, are well known to those of ordinary skill in the art. Still other mutations can be made to the noncoding sequences of a sarcoma-associated antigen gene or cDNA clone to enhance expression of the polypeptide. The activity of variants of sarcoma-associated polypeptides can be tested by cloning the gene encoding the variant sarcoma-associated polypeptide into a bacterial or mammalian expression vector, introducing the vector into an appropriate host cell, expressing the variant sarcoma-associated polypeptide, and testing for a functional capability of the sarcoma-associated polypeptides as disclosed herein. For example, the variant sarcoma-associated polypeptide can be tested for reaction with autologous or allogeneic sera as described in the Examples. Preparation of other variant polypeptides may favor testing of other activities, as will be known to one of ordinary skill in the art.

The skilled artisan will also realize that conservative amino acid substitutions may be made in immunogenic sarcoma-associated polypeptides to provide functionally equivalent variants, or homologs of the foregoing polypeptides, i.e., the variants retain the functional capabilities of the immunogenic sarcoma-associated polypeptides. As used herein, a "conservative amino acid substitution" refers to an amino acid substitution that does not alter the relative charge or size characteristics of the protein in which the amino acid substitution is made. Variants can be prepared according to methods for altering polypeptide sequence known to one of ordinary skill in the art such as are found in references that compile such methods, e.g. Molecular Cloning: A Laboratory Manual, J. Sambrook, et al., eds., Second Edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 1989, or Current Protocols in Molecular Biology, F.M. Ausubel, et al., eds., John Wiley & Sons, Inc., New York. Exemplary functionally equivalent variants or homologs of the sarcoma-associated polypeptides include conservative amino acid substitutions of in the amino acid sequences of proteins disclosed herein. Conservative substitutions of amino acids include

substitutions made amongst amino acids within the following groups: (a) M, I, L, V; (b) F, Y, W; (c) K, R, H; (d) A, G; (e) S, T; (f) Q, N; and (g) E, D. Therefore, one can make conservative amino acid substitutions to the amino acid sequence of the sarcoma-associated antigens disclosed herein and retain the specific antibody-binding characteristics of the antigens.

Likewise, upon determining that a peptide derived from a sarcoma-associated polypeptide is presented by an MHC molecule and recognized by antibodies or T lymphocytes (e.g., helper T cells or CTLs), one can make conservative amino acid substitutions to the amino acid sequence of the peptide, particularly at residues which are thought not to be direct contact points with the MHC molecule. For example, methods for identifying functional variants of HLA class II binding peptides are provided in a published PCT application of Strominger and Wucherpfennig (PCT/US96/03182). Peptides bearing one or more amino acid substitutions also can be tested for concordance with known HLA/MHC motifs prior to synthesis using, e.g. the computer program described by D'Amaro and Drijfhout (D'Amaro et al., *Human Immunol.* 43:13-18, 1995; Drijfhout et al., *Human Immunol.* 43:1-12, 1995). The substituted peptides can then be tested for binding to the MHC molecule and recognition by antibodies or T lymphocytes when bound to MHC. These variants can be tested for improved stability and are useful, *inter alia*, in vaccine compositions.

Conservative amino-acid substitutions in the amino acid sequence of sarcoma-associated polypeptides to produce functionally equivalent variants of sarcoma-associated polypeptides typically are made by alteration of a nucleic acid encoding a sarcoma-associated polypeptide. Such substitutions can be made by a variety of methods known to one of ordinary skill in the art. For example, amino acid substitutions may be made by PCR-directed mutation, site-directed mutagenesis according to the method of Kunkel (Kunkel, Proc. Nat. Acad. Sci. U.S.A. 82: 488-492, 1985), or by chemical synthesis of a gene encoding a sarcoma-associated polypeptide. Where amino acid substitutions are made to a small unique fragment of a sarcoma-associated polypeptide, such as an antigenic epitope recognized by autologous or allogeneic sera or T lymphocytes, the substitutions can be made by directly synthesizing the peptide. The activity of functionally equivalent variants of sarcoma-associated polypeptides can be tested by cloning the gene encoding the altered sarcoma-associated polypeptide into a bacterial or mammalian expression vector, introducing the vector into an appropriate host cell, expressing the altered polypeptide, and testing for a

functional capability of the sarcoma-associated polypeptides as disclosed herein. Peptides that are chemically synthesized can be tested directly for function, e.g., for binding to antisera recognizing associated antigens.

The invention as described herein has a number of uses, some of which are described elsewhere herein. In one aspect of the invention a method of identifying cancer-associated antigens is provided. Novel cancer-associated antigens can be identified by obtaining a biological sample from a subject, determining the reactivity of the biological sample with one or more known cancer-associated antigens, and subsequently using the reactive biological sample to screen an expression library to identify novel cancer-associated antigens as well as proteins previously known but not previously associated with cancer.

As used herein, a "subject" is preferably a human, non-human primate, cow, horse, pig, sheep, goat, dog, cat or rodent. In all embodiments, human subjects are preferred. In some embodiments, the subject is suspected of having cancer or has been diagnosed with cancer. Cancers in which the sarcoma-associated nucleic acid or polypeptide are differentially expressed include sarcoma.

As used herein, a biological sample includes, but is not limited to: tissue, cells, or body fluid (e.g. serum, blood, lymph node fluid, etc.). The fluid sample may include cells and/or fluid. The tissue and cells may be obtained from a subject or may be grown in culture (e.g. from a cell line). As used herein, a biological sample is body fluid, tissue or cells obtained from a subject using methods well-known to those of ordinary skill in the related medical arts.

The invention in another aspect permits the isolation of the cancer-associated antigens described herein. A variety of methodologies well-known to the skilled practitioner can be utilized to obtain isolated cancer-associated antigens. The proteins may be purified from cells which naturally produce the protein by chromatographic means or immunological recognition. Alternatively, an expression vector may be introduced into cells to cause production of the protein. In another method, mRNA transcripts may be microinjected or otherwise introduced into cells to cause production of the encoded protein. Translation of mRNA in cell-free extracts such as the reticulocyte lysate system also may be used to produce the protein. Those skilled in the art also can readily follow known methods for isolating cancer-associated antigens. These include, but are not limited to, chromatographic techniques such as immunochromatography, HPLC, size-exclusion chromatography, ion-exchange chromatography, and immune-affinity chromatography.

The invention also involves diagnosing or monitoring cancer in subjects by determining the presence of an immune response to one or more sarcoma-associated antigens of the invention. In preferred embodiments, this determination is performed by assaying a bodily fluid obtained from the subject, preferably serum, blood, or lymph node fluid for the presence of antibodies against the sarcoma-associated antigens described herein. This determination may also be performed by assaying a tissue or cells from the subject for the presence of one or more sarcoma-associated antigens (or nucleic acid molecules that encode these antigens) described herein. In another embodiment, the presence of antibodies against at least one additional cancer antigen is determined for diagnosis of cancer. The additional antigen may be a sarcoma-associated antigen as described herein or may be some other cancer-associated antigen. This determination may also be performed by assaying a tissue or cells from the subject for the presence of the sarcoma-associated antigens described herein.

Measurement of the immune response against one of the sarcoma-associated antigens over time by sequential determinations permits monitoring of the disease and/or the effects of a course of treatment. For example, a sample, such as serum, blood, or lymph node fluid, may be obtained from a subject, tested for an immune response to one of the sarcoma-associated antigens, and at a second, subsequent time, another sample, may be obtained from the subject and similarly tested. The results of the first and second (or subsequent) tests can be compared as a measure of the onset, regression or progression of cancer, or, if cancer treatment was undertaken during the interval between obtaining the samples, the effectiveness of the treatment may be evaluated by comparing the results of the two tests. In preferred embodiments the sarcoma-associated antigens are bound to a substrate. In other preferred embodiments the immune response of the biological sample to the sarcoma-associated antigens is determined with ELISA. Other methods will be apparent to one of skill in the art.

Diagnostic methods of the invention also involve determining the aberrant expression of one or more of the sarcoma-associated antigens described herein or the nucleic acid molecules that encode them. Such determinations can be carried out via any standard nucleic acid assay, including the polymerase chain reaction or assaying with hybridization probes, which may be labeled, or by assaying biological samples with binding partners (e.g., antibodies) for sarcoma-associated antigens.

The diagnostic methods of the invention can be used to detect the presence of a disorder associated with aberrant expression of a sarcoma-associated molecule, as well as to

assess the progression and/or regression of the disorder such as in response to treatment (e.g., chemotherapy, radiation). According to this aspect of the invention, the method for diagnosing a disorder characterized by aberrant expression of a sarcoma-associated molecule involve: detecting expression of a sarcoma-associated molecule in a first biological sample obtained from a subject, wherein differential expression of the sarcoma-associated molecule compared to a control sample indicates that the subject has a disorder characterized by aberrant expression of a sarcoma-associated molecule, such as cancer.

As used herein, "aberrant expression" of a sarcoma-associated antigen is intended to include any expression that is statistically significant from the expected amount of expression. For example, expression of a sarcoma-associated molecule (i.e., the sarcoma-associated antigen or the nucleic acid molecules that encode it) in a tissue that is not expected to express the sarcoma-associated molecule would be included in the definition of "aberrant expression". Likewise, expression of the sarcoma-associated molecule that is determined to be expressed at a significantly higher or lower level than expected is also included. Therefore, a determination of the level of expression of one or more of the sarcoma-associated antigens and/or the nucleic acids that encode them is diagnostic of cancer if the level of expression is above a baseline level determined for that tissue type. The baseline level of expression can be determined using standard methods known to those of skill in the art. Such methods include, for example, assaying a number of histologically normal tissue samples from subjects that are clinically normal (i.e. do not have clinical signs of cancer in that tissue type) and determining the mean level of expression for the samples.

The level of expression of the nucleic acid molecules of the invention or the antigens they encode can indicate cancer in the tissue when the level of expression is significantly more in the tissue than in a control sample. In some embodiments, a level of expression in the tissues that is at least about 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100%, 150%, 200%, 250%, 300%, 400 %, or 500% more than the level of expression in the control tissue indicates cancer in the tissue.

As used herein the term "control" means predetermined values, and also means samples of materials tested in parallel with the experimental materials. Examples include samples from control populations or control samples generated through manufacture to be tested in parallel with the experimental samples.

As used herein the term "control" includes positive and negative controls which may be a predetermined value that can take a variety of forms. The control(s) can be a single cut-

off value, such as a median or mean, or can be established based upon comparative groups, such as in groups having normal amounts of sarcoma-associated molecules of the invention and groups having abnormal amounts of sarcoma-associated molecules of the invention. Another example of a comparative group is a group having a particular disease, condition and/or symptoms and a group without the disease, condition and/or symptoms. Another comparative group is a group with a family history of a particular disease and a group without such a family history of the particular disease. The predetermined control value can be arranged, for example, where a tested population is divided equally (or unequally) into groups, such as a low-risk group, a medium-risk group and a high-risk group or into quadrants or quintiles, the lowest quadrant or quintile being individuals with the lowest risk or lowest expression levels of a sarcoma-associated molecule of the invention that is up-regulated in cancer and the highest quadrant or quintile being individuals with the highest risk or highest expression levels of a sarcoma-associated molecule of the invention that is up-regulated in cancer.

The predetermined value of a control will depend upon the particular population selected. For example, an apparently healthy population will have a different "normal" sarcoma-associated molecule expression level range than will a population which is known to have a condition characterized by aberrant expression of the sarcoma-associated molecule. Accordingly, the predetermined value selected may take into account the category in which an individual falls. Appropriate ranges and categories can be selected with no more than routine experimentation by those of ordinary skill in the art. Typically the control will be based on apparently healthy individuals in an appropriate age bracket. As used herein, the term "increased expression" means a higher level of expression relative to a selected control.

The invention involves in some aspects diagnosing or monitoring cancer by determining the level of expression of one or more sarcoma-associated nucleic acid molecules and/or determining the level of expression of one or more sarcoma-associated polypeptides they encode. In some important embodiments, this determination is performed by assaying a tissue sample from a subject for the level of expression of one or more sarcoma-associated nucleic acid molecules or for the level of expression of one or more sarcoma-associated polypeptides encoded by the nucleic acid molecules of the invention.

The expression of the molecules of the invention may be determined using routine methods known to those of ordinary skill in the art. These methods include, but are not limited to: direct RNA amplification, reverse transcription of RNA to cDNA, real-time RT-

PCR, amplification of cDNA, hybridization, and immunologically based assay methods, which include, but are not limited to immunohistochemistry, antibody sandwich capture assay, ELISA, and enzyme-linked immunospot assay (EliSpot assay). For example, the determination of the presence of level of nucleic acid molecules of the invention in a subject or tissue can be carried out via any standard nucleic acid determination assay, including the polymerase chain reaction, or assaying with labeled hybridization probes. Such hybridization methods include, but are not limited to microarray techniques.

These methods of determining the presence and/or level of the molecules of the invention in cells and tissues may include use of labels to monitor the presence of the molecules of the invention. Such labels may include, but are not limited to radiolabels or chemiluminescent labels, which may be utilized to determine whether a molecule of the invention is expressed in a cell or tissue, and to determine the level of expression in the cell or tissue. For example, a fluorescently labeled or radiolabeled antibody that selectively binds to a polypeptide of the invention may be contacted with a tissue or cell to visualize the polypeptide in vitro or in vivo. These and other in vitro and in vivo imaging methods for determining the presence of the nucleic acid and polypeptide molecules of the invention are well known to those of ordinary skill in the art.

The invention, therefore, also involves the use of agents such as polypeptides that bind to sarcoma-associated antigens. Such agents can be used in methods of the invention including the diagnosis and/or treatment of cancer. Such binding agents can be used, for example, in screening assays to detect the presence or absence of sarcoma-associated antigens and can be used in quantitative binding assays to determine levels of expression in biological samples and cells. Such agents also may be used to inhibit the native activity of the sarcoma-associated polypeptides, for example, by binding to such polypeptides.

According to this aspect, the binding polypeptides bind to an isolated nucleic acid or protein of the invention, including unique fragments thereof. Preferably, the binding polypeptides bind to a sarcoma-associated polypeptide, or a unique fragment thereof.

In preferred embodiments, the binding polypeptide is an antibody or antibody fragment, more preferably, an Fab or $F(ab)_2$ fragment of an antibody. Typically, the fragment includes a CDR3 region that is selective for the sarcoma-associated antigen. Any of the various types of antibodies can be used for this purpose, including polyclonal antibodies, monoclonal antibodies, humanized antibodies, and chimeric antibodies.

Thus, the invention provides agents which bind to sarcoma-associated antigens encoded by sarcoma-associated nucleic acid molecules of the invention, and in certain embodiments preferably to unique fragments of the sarcoma-associated polypeptides. Such binding partners can be used in screening assays to detect the presence or absence of a sarcoma-associated antigen and in purification protocols to isolate such sarcoma-associated antigens. Likewise, such binding partners can be used to selectively target drugs, toxins or other molecules (including detectable diagnostic molecules) to cells which express sarcoma-associated antigens. In this manner, for example, cells present in solid or non-solid tumors which express sarcoma-associated proteins can be treated with cytotoxic compounds that are selective for the sarcoma-associated molecules (nucleic acids and/or antigens). Such binding agents also can be used to inhibit the native activity of the sarcoma-associated antigen, for example, to further characterize the functions of these molecules.

The antibodies of the present invention thus are prepared by any of a variety of methods, including administering a protein, fragments of a protein, cells expressing the protein or fragments thereof and the like to an animal to induce polyclonal antibodies. The present invention also provides methods of producing monoclonal antibodies to the sarcoma-associated molecules of the invention described herein. The production of monoclonal antibodies is according to techniques well known in the art. As detailed herein, such antibodies may be used for example to identify tissues expressing protein or to purify protein. Antibodies also may be coupled to specific labeling agents or imaging agents, including, but not limited to a molecule preferably selected from the group consisting of fluorescent, enzyme, radioactive, metallic, biotin, chemiluminescent, bioluminescent, chromophore, or colored, etc. In some aspects of the invention, a label may be a combination of the foregoing molecule types.

Significantly, as is well-known in the art, only a small portion of an antibody molecule, the paratope, is involved in the binding of the antibody to its epitope (see, in general, Clark, W.R. (1986) The Experimental Foundations of Modern Immunology Wiley & Sons, Inc., New York; Roitt, I. (1991) Essential Immunology, 7th Ed., Blackwell Scientific Publications, Oxford). The pFc' and Fc regions, for example, are effectors of the complement cascade but are not involved in antigen binding. An antibody from which the pFc' region has been enzymatically cleaved, or which has been produced without the pFc' region, designated an F(ab')2 fragment, retains both of the antigen binding sites of an intact antibody. Similarly, an antibody from which the Fc region has been enzymatically cleaved, or which has been

produced without the Fc region, designated an Fab fragment, retains one of the antigen binding sites of an intact antibody molecule. Fab fragments consist of a covalently bound antibody light chain and a portion of the antibody heavy chain denoted Fd. The Fd fragments are the major determinant of antibody specificity (a single Fd fragment may be associated with up to ten different light chains without altering antibody specificity) and Fd fragments retain epitope-binding ability in isolation.

Within the antigen-binding portion of an antibody, as is well-known in the art, there are complementarity determining regions (CDRs), which directly interact with the epitope of the antigen, and framework regions (FRs), which maintain the tertiary structure of the paratope (see, in general, Clark, 1986; Roitt, 1991). In both the heavy chain Fd fragment and the light chain of IgG immunoglobulins, there are four framework regions (FR1 through FR4) separated respectively by three complementarity determining regions (CDR1 through CDR3). The CDRs, and in particular the CDR3 regions, and more particularly the heavy chain CDR3, are largely responsible for antibody specificity.

It is now well-established in the art that the non-CDR regions of a mammalian antibody may be replaced with similar regions of nonspecific or heterospecific antibodies while retaining the epitopic specificity of the original antibody. This is most clearly manifested in the development and use of "humanized" antibodies in which non-human CDRs are covalently joined to human FR and/or Fc/pFc' regions to produce a functional antibody. See, e.g., U.S. patents 4,816,567, 5,225,539, 5,585,089, 5,693,762, and 5,859,205.

Fully human monoclonal antibodies also can be prepared by immunizing mice transgenic for large portions of human immunoglobulin heavy and light chain loci. Following immunization of these mice (e.g., XenoMouse (Abgenix), HuMAb mice (Medarex/GenPharm)), monoclonal antibodies can be prepared according to standard hybridoma technology. These monoclonal antibodies will have human immunoglobulin amino acid sequences and therefore will not provoke human anti-mouse antibody (HAMA) responses when administered to humans.

Thus, as will be apparent to one of ordinary skill in the art, the present invention also provides for F(ab')2, Fab, Fv, and Fd fragments; chimeric antibodies in which the Fc and/or FR and/or CDR1 and/or CDR2 and/or light chain CDR3 regions have been replaced by homologous human or non-human sequences; chimeric F(ab')2 fragment antibodies in which the FR and/or CDR1 and/or CDR2 and/or light chain CDR3 regions have been replaced by homologous human or non-human sequences; chimeric Fab fragment antibodies in which the

FR and/or CDR1 and/or CDR2 and/or light chain CDR3 regions have been replaced by homologous human or non-human sequences; and chimeric Fd fragment antibodies in which the FR and/or CDR1 and/or CDR2 regions have been replaced by homologous human or non-human sequences. The present invention also includes so-called single chain antibodies.

Thus, the invention involves polypeptides of numerous size and type that bind specifically to sarcoma-associated antigens. These polypeptides may be derived also from sources other than antibody technology. For example, such polypeptide binding agents can be provided by degenerate peptide libraries which can be readily prepared in solution, in immobilized form or as phage display libraries. Combinatorial libraries also can be synthesized of peptides containing one or more amino acids. Libraries further can be synthesized of peptides and non-peptide synthetic moieties.

The sarcoma-associated antigens of the invention can be used to screen peptide libraries, including phage display libraries, to identify and select peptide binding partners of the sarcoma-assoc iated antigens of the invention. Such molecules can be used, as described, for screening assays, for diagnostic assays, for purification protocols or for targeting drugs, toxins and/or labeling agents (e.g., radioisotopes, fluorescent molecules, etc.) to cells which express sarcoma-associated molecules such as cancer cells which have aberrant sarcoma-associated expression.

Phage display can be particularly effective in identifying binding peptides useful according to the invention. Briefly, one prepares a phage library (using e.g. m13, fd, or lambda phage), displaying inserts from 4 to about 80 amino acid residues using conventional procedures. The inserts may represent, for example, a completely degenerate or biased array. One then can select phage-bearing inserts which bind to the sarcoma-associated antigen. This process can be repeated through several cycles of reselection of phage that bind to the sarcoma-associated polypeptide. Repeated rounds lead to enrichment of phage bearing particular sequences. DNA sequence analysis can be conducted to identify the sequences of the expressed polypeptides. The minimal linear portion of the sequence that binds to the sarcoma-associated polypeptide can be determined. One can repeat the procedure using a biased library containing inserts containing part or all of the minimal linear portion plus one or more additional degenerate residues upstream or downstream thereof. Yeast two-hybrid screening methods also may be used to identify polypeptides that bind to the sarcoma-associated antigens.

As detailed herein, the foregoing antibodies and other binding molecules may be used to identify tissues with normal or aberrant expression of a sarcoma-associated antigen. Antibodies also may be coupled to specific diagnostic labeling agents for imaging of cells and tissues with normal or aberrant sarcoma-associated antigen expression or to therapeutically useful agents according to standard coupling procedures. As used herein, "therapeutically useful agents" include any therapeutic molecule which desirably is targeted selectively to a cell or tissue selectively with an aberrant sarcoma-associated expression.

Diagnostic agents for *in vivo* use include, but are not limited to, barium sulfate, iocetamic acid, iopanoic acid, ipodate calcium, diatrizoate sodium, diatrizoate meglumine, metrizamide, tyropanoate sodium and radiodiagnostics including positron emitters such as fluorine-18 and carbon-11, gamma emitters such as iodine-123, technitium-99, iodine-131 and indium-111, and nuclides for nuclear magnetic resonance such as fluorine and gadolinium. Other diagnostic agents useful in the invention will be apparent to one of ordinary skill in the art.

The antibodies of the present invention can also be used to therapeutically target sarcoma-associated antigens. In a preferred embodiment, antibodies can be used to target antigens expressed on the cell surface, such as NY-SAR-35. These antibodies can be linked not only to a detectable marker but also an antitumor agent or an immunomodulator. Antitumor agents can include cytotoxic agents and agents that act on tumor neovasculature. Detectable markers include, for example, radioactive or fluorescent markers. Cytotoxic agents include cytotoxic radionuclides, chemical toxins and protein toxins.

The cytotoxic radionuclide or radiotherapeutic isotope preferably is an alpha-emitting isotope such as $^{225}Ac$, $^{211}At$, $^{212}Bi$, $^{213}Bi$, $^{212}Pb$, $^{224}Ra$ or $^{223}Ra$. Alternatively, the cytotoxic radionuclide may a beta-emitting isotope such as $^{186}Rh$, $^{188}Rh$, $^{177}Lu$, $^{90}Y$, $^{131}I$, $^{67}Cu$, $^{64}Cu$, $^{153}Sm$ or $^{166}Ho$. Further, the cytotoxic radionuclide may emit Auger and low energy electrons and include the isotopes $^{125}I$, $^{123}I$ or $^{77}Br$.

Suitable chemical toxins or chemotherapeutic agents include members of the enediyne family of molecules, such as calicheamicin and esperamicin. Chemical toxins can also be taken from the group consisting of methotrexate, doxorubicin, melphalan, chlorambucil, ARA-C, vindesine, mitomycin C, cis-platinum, etoposide, bleomycin and 5-fluorouracil. Other antineoplastic agents that may be conjugated to the anti-PSMA antibodies of the present invention include dolastatins (U.S. Patent Nos. 6,034,065 and 6,239,104) and derivatives thereof. Of particular interest is dolastatin 10 (dolavaline-valine-dolaisoleuine-

dolaproine-dolaphenine) and the derivatives auristatin PHE (dolavaline-valine-dolaisoleuine-dolaproine-phenylalanine-methyl ester) (Pettit, G.R. et al., *Anticancer Drug Des.* 13(4):243-277, 1998; Woyke, T. et al., *Antimicrob. Agents Chemother.* 45(12):3580-3584, 2001), and aurastatin E and the like. Toxins that are less preferred in the compositions and methods of the invention include poisonous lectins, plant toxins such as ricin, abrin, modeccin, botulina and diphtheria toxins. Of course, combinations of the various toxins could also be coupled to one antibody molecule thereby accommodating variable cytotoxicity. Other chemotherapeutic agents are known to those skilled in the art.

Agents that act on the tumor vasculature can include tubulin-binding agents such as combrestatin A4 (Griggs et al., *Lancet Oncol.* 2:82, 2001), angiostatin and endostatin (reviewed in Rosen, *Oncologist* 5:20, 2000, incorporated by reference herein) and interferon inducible protein 10 (U.S. Patent No. 5,994,292). A number of antiangiogenic agents currently in clinical trials are also contemplated. Agents currently in clinical trials include: 2ME2, Angiostatin, Angiozyme, Anti-VEGF RhuMAb, Apra (CT-2584), Avicine, Benefin, BMS275291, Carboxyamidotriazole, CC4047, CC5013, CC7085, CDC801, CGP-41251 (PKC 412), CM101, Combretastatin A-4 Prodrug, EMD 121974, Endostatin, Flavopiridol, Genistein (GCP), Green Tea Extract, IM-862, ImmTher, Interferon alpha, Interleukin-12, Iressa (ZD1839), Marimastat, Metastat (Col-3), Neovastat , Octreotide, Paclitaxel, Penicillamine, Photofrin, Photopoint, PI-88, Prinomastat (AG-3340), PTK787 (ZK22584), RO317453, Solimastat, Squalamine, SU 101, SU 5416, SU-6668, Suradista (FCE 26644), Suramin (Metaret), Tetrathiomolybdate, Thalidomide, TNP-470 and Vitaxin. additional antiangiogenic agents are described by Kerbel, J. Clin. Oncol. 19(18s):45s-51s, 2001, which is incorporated by reference herein. Immunomodulators suitable for conjugation to the antibodies include α-interferon, γ-interferon, and tumor necrosis factor alpha (TNFα).

The coupling of one or more toxin molecules to the antibody is envisioned to include many chemical mechanisms, for instance covalent binding, affinity binding, intercalation, coordinate binding, and complexation. The toxic compounds used to prepare the immunotoxins are attached to the antibodies or antigen-binding fragments thereof by standard protocols known in the art.

In other aspects of the invention, the sarcoma-associated molecules and the antibodies and other binding molecules, as described herein, can be used for the treatment of disorders. When "disorder" is used herein, it refers to any pathological condition where the sarcoma-associated antigens are aberrantly expressed. An example of such a disorder is cancer, with

sarcoma as a particular example. For human cancers, additional particular examples include synovial sarcoma, liposarcoma, neurosarcoma, chondrosarcoma, fibrosarcoma, Ewing sarcoma, leimyosarcoma, osteosarcoma, rhabdomyosarcoma, malignant fibrous histocytoma, DFSP, leukemia, lymphoma, gastric cancer, glioma, bladder cancer, breast cancer, ovarian cancer, renal cancer, lung cancer, colon cancer, prostate cancer, esophageal cancer, melanoma and hepatoma.

Conventional treatment for cancer may include, but is not limited to: surgical intervention, chemotherapy, radiotherapy, and adjuvant systemic therapies. In one aspect of the invention, treatment may include administering binding polypeptides such as antibodies that specifically bind to the sarcoma-associated antigen. These binding polypeptides can be optionally linked to one or more detectable markers, antitumor agents or immunomodulators as described above.

Cancer treatment, in another aspect of the invention may include administering an antisense molecules or RNAi molecules to reduce expression level and/or function level of sarcoma-associated polypeptides of the invention in the subject in cancers where a sarcoma-associated molecule is up-regulated. The use of RNA interference or "RNAi" involves the use of double-stranded RNA (dsRNA) to block gene expression. (see: Sui, G, et al, Proc Natl. Acad. Sci U.S.A. 99:5515-5520,2002). Methods of applying RNAi strategies in embodiments of the invention would be understood by one of ordinary skill in the art.

Sarcoma-associated polypeptides as described herein, can also be used in one aspect of the invention to induce or enhance an immune response. Some therapeutic approaches based upon the disclosure are premised on a response by a subject's immune system, leading to lysis of antigen presenting cells, such as cancer cells which present one or more sarcoma-associated antigens of the invention. One such approach is the administration of autologous CTLs specific to a sarcoma-associated antigen/MHC complex to a subject with abnormal cells of the phenotype at issue. It is within the ability of one of ordinary skill in the art to develop such CTLs in vitro. An example of a method for T cell differentiation is presented in International Application number PCT/US96/05607. Generally, a sample of cells taken from a subject, such as blood cells, are contacted with a cell presenting the complex and capable of provoking CTLs to proliferate. The target cell can be a transfectant, such as a COS cell. These transfectants present the desired complex of their surface and, when combined with a CTL of interest, stimulate its proliferation. COS cells are widely available, as are other

suitable host cells. Specific production of CTL clones is well known in the art. The clonally expanded autologous CTLs then are administered to the subject.

Another method for selecting antigen-specific CTL clones has recently been described (Altman et al., Science 274:94-96, 1996; Dunbar et al., Curr. Biol. 8:413-416, 1998), in which fluorogenic tetramers of MHC class I molecule/peptide complexes are used to detect specific CTL clones. Briefly, soluble MHC class I molecules are folded in vitro in the presence of $\beta_2$-microglobulin and a peptide antigen which binds the class I molecule. After purification, the MHC/peptide complex is purified and labeled with biotin. Tetramers are formed by mixing the biotinylated peptide-MHC complex with labeled avidin (e.g. phycoerythrin) at a molar ratio or 4:1. Tetramers are then contacted with a source of CTLs such as peripheral blood or lymph node. The tetramers bind CTLs which recognize the peptide antigen/MHC class I complex. Cells bound by the tetramers can be sorted by fluorescence activated cell sorting to isolate the reactive CTLs. The isolated CTLs then can be expanded in vitro for use as described herein.

To detail a therapeutic methodology, referred to as adoptive transfer (Greenberg, J. Immunol. 136(5): 1917, 1986; Riddel et al., Science 257: 238, 1992; Lynch et al, Eur. J. Immunol. 21: 1403-1410,1991; Kast et al., Cell 59: 603-614, 1989), cells presenting the desired complex (e.g., dendritic cells) are combined with CTLs leading to proliferation of the CTLs specific thereto. The proliferated CTLs are then administered to a subject with a cellular abnormality which is characterized by certain of the abnormal cells presenting the particular complex. The CTLs then lyse the abnormal cells, thereby achieving the desired therapeutic goal.

The foregoing therapy assumes that at least some of the subject's abnormal cells present the relevant HLA/cancer associated antigen complex. This can be determined very easily, as the art is very familiar with methods for identifying cells which present a particular HLA molecule, as well as how to identify cells expressing DNA of the pertinent sequences, in this case a sarcoma-associated antigen sequence. Once cells presenting the relevant complex are identified via the foregoing screening methodology, they can be combined with a sample from a patient, where the sample contains CTLs. If the complex presenting cells are lysed by the mixed CTL sample, then it can be assumed that a sarcoma-associated antigen is being presented, and the subject is an appropriate candidate for the therapeutic approaches set forth supra.

Adoptive transfer is not the only form of therapy that is available in accordance with the invention. CTLs can also be provoked in vivo, using a number of approaches. One approach is the use of non-proliferative cells expressing the complex. The cells used in this approach may be those that normally express the complex, such as irradiated tumor cells or cells transfected with one or both of the genes necessary for presentation of the complex (i.e. the antigenic peptide and the presenting MHC molecule). Chen et al. (Proc. Natl. Acad. Sci. USA 88: 110-114,1991) exemplifies this approach, showing the use of transfected cells expressing HPV E7 peptides in a therapeutic regime. Various cell types may be used. Similarly, vectors carrying one or both of the genes of interest may be used. Viral or bacterial vectors are especially preferred. For example, nucleic acids which encode a sarcoma-associated polypeptide may be operably linked to promoter and enhancer sequences which direct expression of the sarcoma-associated antigen polypeptide in certain tissues or cell types. The nucleic acid may be incorporated into an expression vector.

Expression vectors may be unmodified extrachromosomal nucleic acids, plasmids or viral genomes constructed or modified to enable insertion of exogenous nucleic acids, such as those encoding sarcoma-associated antigen, as described elsewhere herein. Nucleic acids encoding a sarcoma-associated antigen also may be inserted into a retroviral genome, thereby facilitating integration of the nucleic acid into the genome of the target tissue or cell type. In these systems, the gene of interest is carried by a microorganism, e.g., a Vaccinia virus, pox virus, herpes simplex virus, retrovirus or adenovirus, and the materials de facto "infect" host cells. The cells which result present the complex of interest, and are recognized by autologous CTLs, which then proliferate.

A similar effect can be achieved by combining the sarcoma-associated polypeptide or a stimulatory fragment thereof with an adjuvant to facilitate incorporation into antigen presenting cells in vivo. The sarcoma-associated polypeptide is processed to yield the peptide partner of the MHC molecule while a sarcoma-associated fragment may be presented without the need for further processing. Generally, subjects can receive an intradermal injection of an effective amount of the sarcoma-associated antigen. Initial doses can be followed by booster doses, following immunization protocols standard in the art. Preferred sarcoma-associated antigens include those found to react with allogeneic cancer antisera, shown in the examples below.

The invention involves the use of various materials disclosed herein to "immunize" subjects or as "vaccines". As used herein, "immunization" or "vaccination" means increasing

or activating an immune response against an antigen. It does not require elimination or eradication of a condition but rather contemplates the clinically favorable enhancement of an immune response toward an antigen. Generally accepted animal models, can be used for testing of immunization against cancer using a sarcoma-associated nucleic acid. For example, human cancer cells can be introduced into a mouse to create a tumor, and one or more sarcoma-associated nucleic acids can be delivered by the methods described herein. The effect on the cancer cells (e.g., reduction of tumor size) can be assessed as a measure of the effectiveness of the sarcoma-associated nucleic acid immunization. Of course, testing of the foregoing animal model using more conventional methods for immunization include the administration of one or more sarcoma-associated polypeptides or fragments derived therefrom, optionally combined with one or more adjuvants and/or cytokines to boost the immune response.

Methods for immunization, including formulation of a vaccine composition and selection of doses, route of administration and the schedule of administration (e.g. primary and one or more booster doses), are well known in the art. The tests also can be performed in humans, where the end point is to test for the presence of enhanced levels of circulating CTLs against cells bearing the antigen, to test for levels of circulating antibodies against the antigen, to test for the presence of cells expressing the antigen and so forth.

As part of the immunization compositions, one or more sarcoma-associated polypeptides or immunogenic fragments thereof are administered with one or more adjuvants to induce an immune response or to increase an immune response. An adjuvant is a substance incorporated into or administered with antigen which potentiates the immune response. Adjuvants may enhance the immunological response by providing a reservoir of antigen (extracellularly or within macrophages), activating macrophages and stimulating specific sets of lymphocytes. Adjuvants of many kinds are well known in the art. Specific examples of adjuvants include monophosphoryl lipid A (MPL, SmithKline Beecham), a congener obtained after purification and acid hydrolysis of Salmonella minnesota Re 595 lipopolysaccharide; saponins including QS21 (SmithKline Beecham), a pure QA-21 saponin purified from Quillja saponaria extract; DQS21, described in PCT application WO96/33739 (SmithKline Beecham); QS-7, QS-17, QS-18, and QS-L1 (So et al., Mol. Cells 7:178-186, 1997); incomplete Freund's adjuvant; complete Freund's adjuvant; montanide; alum; CpG oligonucleotides (see e.g. Kreig et al., Nature 374:546-9, 1995); and various water-in-oil emulsions prepared from biodegradable oils such as squalene and/or tocopherol. Preferably,

the antigens are administered mixed with a combination of DQS21/MPL. The ratio of DQS21 to MPL typically will be about 1:10 to 10:1, preferably about 1:5 to 5:1 and more preferably about 1:1. Typically for human administration, DQS21 and MPL will be present in a vaccine formulation in the range of about 1 μg to about 100 μg. Other adjuvants are known in the art and can be used in the invention (see, e.g. Goding, Monoclonal Antibodies: Principles and Practice, 2nd Ed., 1986). Methods for the preparation of mixtures or emulsions of polypeptide and adjuvant are well known to those of skill in the art of vaccination.

Other agents which stimulate the immune response of the subject can also be administered to the subject. For example, other cytokines are also useful in vaccination protocols as a result of their lymphocyte regulatory properties. Many other cytokines useful for such purposes will be known to one of ordinary skill in the art, including interleukin-12 (IL-12) which has been shown to enhance the protective effects of vaccines (see, e.g., Science 268: 1432-1434, 1995), GM-CSF and IL-18. Thus cytokines can be administered in conjunction with antigens and adjuvants to increase the immune response to the antigens.

There are a number of immune response potentiating compounds that can be used in vaccination protocols. These include costimulatory molecules provided in either protein or nucleic acid form. Such costimulatory molecules include the B7-1 and B7-2 (CD80 and CD86 respectively) molecules which are expressed on dendritic cells (DC) and interact with the CD28 molecule expressed on the T cell. This interaction provides costimulation (signal 2) to an antigen/MHC/TCR stimulated (signal 1) T cell, increasing T cell proliferation and effector function. B7 also interacts with CTLA4 (CD152) on T cells and studies involving CTLA4 and B7 ligands indicate that the B7-CTLA4 interaction can enhance antitumor immunity and CTL proliferation (Zheng P., et al. Proc. Natl. Acad. Sci. USA 95 (11):6284-6289 (1998)).

B7 typically is not expressed on tumor cells so they are not efficient antigen presenting cells (APCs) for T cells. Induction of B7 expression would enable the tumor cells to stimulate more efficiently CTL proliferation and effector function. A combination of B7/IL-6/IL-12 costimulation has been shown to induce IFN-gamma and a Th1 cytokine profile in the T cell population leading to further enhanced T cell activity (Gajewski et al., J. Immunol, 154:5637-5648 (1995)). Tumor cell transfection with B7 has ben discussed in relation to in vitro CTL expansion for adoptive transfer immunotherapy by Wang et al., (J. Immunol., 19:1-8 (1986)). Other delivery mechanisms for the B7 molecule would include

nucleic acid (naked DNA) immunization (Kim J., et al. Nat. Biotechnol., 15:7:641-646 (1997)) and recombinant viruses such as adeno and pox (Wendtner et al., Gene Ther., 4:7:726-735 (1997)). These systems are all amenable to the construction and use of expression cassettes for the coexpression of B7 with other molecules of choice such as the antigens or fragment(s) of antigens discussed herein (including polytopes) or cytokines. These delivery systems can be used for induction of the appropriate molecules in vitro and for in vivo vaccination situations. The use of anti-CD28 antibodies to directly stimulate T cells in vitro and in vivo could also be considered. Similarly, the inducible co-stimulatory molecule ICOS which induces T cell responses to foreign antigen could be modulated, for example, by use of anti-ICOS antibodies (Hutloff et al., Nature 397:263-266, 1999).

Lymphocyte function associated antigen-3 (LFA-3) is expressed on APCs and some tumor cells and interacts with CD2 expressed on T cells. This interaction induces T cell IL-2 and IFN-gamma production and can thus complement but not substitute, the B7/CD28 costimulatory interaction (Parra et al., J. Immunol., 158:637-642 (1997), Fenton et al., J. Immunother., 21:2:95-108 (1998)).

Lymphocyte function associated antigen-1 (LFA-1) is expressed on leukocytes and interacts with ICAM-1 expressed on APCs and some tumor cells. This interaction induces T cell IL-2 and IFN-gamma production and can thus complement but not substitute, the B7/CD28 costimulatory interaction (Fenton et al., J. Immunother., 21:2:95-108 (1998)). LFA-1 is thus a further example of a costimulatory molecule that could be provided in a vaccination protocol in the various ways discussed above for B7.

Complete CTL activation and effector function requires Th cell help through the interaction between the Th cell CD40L (CD40 ligand) molecule and the CD40 molecule expressed by DCs (Ridge et al., Nature, 393:474 (1998), Bennett et al., Nature, 393:478 (1998), Schoenberger et al., Nature, 393:480 (1998)). This mechanism of this costimulatory signal is likely to involve upregulation of B7 and associated IL-6/IL-12 production by the DC (APC). The CD40-CD40L interaction thus complements the signal 1 (antigen/MHC-TCR) and signal 2 (B7-CD28) interactions.

The use of anti-CD40 antibodies to stimulate DC cells directly, would be expected to enhance a response to tumor antigens which are normally encountered outside of an inflammatory context or are presented by non-professional APCs (tumor cells). In these situations Th help and B7 costimulation signals are not provided.

The invention contemplates delivery of nucleic acids, polypeptides or fragments thereof for vaccination. Delivery of polypeptides and fragments thereof can be accomplished according to standard vaccination protocols which are well known in the art. In another embodiment, the delivery of nucleic acid is accomplished by ex vivo methods, i.e. by removing a cell from a subject, genetically engineering the cell to include a sarcoma-associated polypeptide, and reintroducing the engineered cell into the subject. One example of such a procedure is outlined in U.S. Patent 5,399,346 and in exhibits submitted in the file history of that patent, all of which are publicly available documents. In general, it involves introduction in vitro of a functional copy of a gene into a cell(s) of a subject, and returning the genetically engineered cell(s) to the subject. The functional copy of the gene is under operable control of regulatory elements which permit expression of the gene in the genetically engineered cell(s). Numerous transfection and transduction techniques as well as appropriate expression vectors are well known to those of ordinary skill in the art, some of which are described in PCT application WO95/00654. In vivo nucleic acid delivery using vectors such as viruses and targeted liposomes also is contemplated according to the invention.

A virus vector for delivering a nucleic acid encoding a sarcoma-associated polypeptide is selected from the group consisting of adenoviruses, adeno-associated viruses, poxviruses including vaccinia viruses and attenuated poxviruses, Semliki Forest virus, Venezuelan equine encephalitis virus, retroviruses, Sindbis virus, and Ty virus-like particle. Examples of viruses and virus-like particles which have been used to deliver exogenous nucleic acids include: replication-defective adenoviruses (e.g., Xiang et al., Virology 219:220-227, 1996; Eloit et al., J. Virol. 7:5375-5381, 1997; Chengalvala et al., Vaccine 15:335-339, 1997), a modified retrovirus (Townsend et al., J. Virol. 71:3365-3374, 1997), a nonreplicating retrovirus (Irwin et al., J. Virol. 68:5036-5044, 1994), a replication defective Semliki Forest virus (Zhao et al., Proc. Natl. Acad. Sci. USA 92:3009-3013, 1995), canarypox virus and highly attenuated vaccinia virus derivative (Paoletti, Proc. Natl. Acad. Sci. USA 93:11349-11353, 1996), non-replicative vaccinia virus (Moss, Proc. Natl. Acad. Sci. USA 93:11341-11348, 1996), replicative vaccinia virus (Moss, Dev. Biol. Stand. 82:55-63, 1994), Venzuelan equine encephalitis virus (Davis et al., J. Virol. 70:3781-3787, 1996), Sindbis virus (Pugachev et al., Virology 212:587-594, 1995), and Ty virus-like particle (Allsopp et al., Eur. J. Immunol 26:1951-1959, 1996). A preferred virus vector is an adenovirus.

Preferably the foregoing nucleic acid delivery vectors: (1) contain exogenous genetic material that can be transcribed and translated in a mammalian cell and that can induce an immune response in a host, and (2) contain on a surface a ligand that selectively binds to a receptor on the surface of a target cell, such as a mammalian cell, and thereby gains entry to the target cell.

Various techniques may be employed for introducing nucleic acids of the invention into cells, depending on whether the nucleic acids are introduced in vitro or in vivo in a host. Such techniques include transfection of nucleic acid-$CaPO_4$ precipitates, transfection of nucleic acids associated with DEAE, transfection or infection with the foregoing viruses including the nucleic acid of interest, liposome mediated transfection, and the like. For certain uses, it is preferred to target the nucleic acid to particular cells. In such instances, a vehicle used for delivering a nucleic acid of the invention into a cell (e.g., a retrovirus, or other virus; a liposome) can have a targeting molecule attached thereto. For example, a molecule such as an antibody specific for a surface membrane protein on the target cell or a ligand for a receptor on the target cell can be bound to or incorporated within the nucleic acid delivery vehicle. Preferred antibodies include antibodies which selectively bind a sarcoma-associated antigen, alone or as a complex with a MHC molecule. Especially preferred are monoclonal antibodies. Where liposomes are employed to deliver the nucleic acids of the invention, proteins which bind to a surface membrane protein associated with endocytosis may be incorporated into the liposome formulation for targeting and/or to facilitate uptake. Such proteins include capsid proteins or fragments thereof tropic for a particular cell type, antibodies for proteins which undergo internalization in cycling, proteins that target intracellular localization and enhance intracellular half life, and the like. Polymeric delivery systems also have been used successfully to deliver nucleic acids into cells, as is known by those skilled in the art. Such systems even permit oral delivery of nucleic acids.

According to a further aspect of the invention, compositions containing the nucleic acid molecules, proteins, and binding polypeptides of the invention are provided. The compositions contain any of the foregoing therapeutic agents in an optional pharmaceutically acceptable carrier. Thus, in a related aspect, the invention provides a method for forming a medicament that involves placing a therapeutically effective amount of the therapeutic agent in the pharmaceutically acceptable carrier to form one or more doses. The effectiveness of treatment or prevention methods of the invention can be determined using standard diagnostic methods described herein.

When administered, the therapeutic compositions of the present invention are administered in pharmaceutically acceptable preparations. Such preparations may routinely contain pharmaceutically acceptable concentrations of salt, buffering agents, preservatives, compatible carriers, supplementary immune potentiating agents such as adjuvants and cytokines, and optionally other therapeutic agents.

As used herein, the term "pharmaceutically acceptable" means a non-toxic material that does not interfere with the effectiveness of the biological activity of the active ingredients. The term "physiologically acceptable" refers to a non-toxic material that is compatible with a biological system such as a cell, cell culture, tissue, or organism. The characteristics of the carrier will depend on the route of administration. Physiologically and pharmaceutically acceptable carriers include diluents, fillers, salts, buffers, stabilizers, solubilizers, and other materials which are well known in the art. The term "carrier" denotes an organic or inorganic ingredient, natural or synthetic, with which the active ingredient is combined to facilitate the application. The components of the pharmaceutical compositions also are capable of being co-mingled with the molecules of the present invention, and with each other, in a manner such that there is no interaction which would substantially impair the desired pharmaceutical efficacy.

The therapeutics of the invention can be administered by any conventional route, including injection or by gradual infusion over time. The administration may, for example, be oral, intravenous, intratumoral, intraperitoneal, intramuscular, intracavity, subcutaneous, or transdermal. When antibodies are used therapeutically, a preferred route of administration is by pulmonary aerosol. Techniques for preparing aerosol delivery systems containing antibodies are well known to those of skill in the art. Generally, such systems should utilize components which will not significantly impair the biological properties of the antibodies, such as the paratope binding capacity (see, for example, Sciarra and Cutie, "Aerosols," in Remington's Pharmaceutical Sciences, 18th edition, 1990, pp 1694-1712). Those of skill in the art can readily determine the various parameters and conditions for producing antibody aerosols without undue experimentation. When using antisense preparations of the invention, slow intravenous administration is preferred.

The compositions of the invention are administered in effective amounts. An "effective amount" is that amount of a sarcoma-associated polypeptide composition that alone, or together with further doses, produces the desired response, e.g. increases an immune response to the sarcoma-associated polypeptide. In the case of treating a particular disease or

condition characterized by expression of one or more sarcoma-associated polypeptides, such as cancer, the desired response is inhibiting the progression of the disease. This may involve only slowing the progression of the disease temporarily, although more preferably, it involves halting the progression of the disease permanently. This can be monitored by routine methods or can be monitored according to diagnostic methods of the invention discussed herein. The desired response to treatment of the disease or condition also can be delaying the onset or even preventing the onset of the disease or condition.

Such amounts will depend, of course, on the particular condition being treated, the severity of the condition, the individual patient parameters including age, physical condition, size and weight, the duration of the treatment, the nature of concurrent therapy (if any), the specific route of administration and like factors within the knowledge and expertise of the health practitioner. These factors are well known to those of ordinary skill in the art and can be addressed with no more than routine experimentation. It is generally preferred that a maximum dose of the individual components or combinations thereof be used, that is, the highest safe dose according to sound medical judgment. It will be understood by those of ordinary skill in the art, however, that a patient may insist upon a lower dose or tolerable dose for medical reasons, psychological reasons or for virtually any other reasons.

The pharmaceutical compositions used in the foregoing methods preferably are sterile and contain an effective amount of sarcoma-associated polypeptide or nucleic acid encoding sarcoma-associated polypeptide for producing the desired response in a unit of weight or volume suitable for administration to a patient. The response can, for example, be measured by determining the immune response following administration of the sarcoma-associated polypeptide composition via a reporter system by measuring downstream effects such as gene expression, or by measuring the physiological effects of the sarcoma-associated polypeptide composition, such as regression of a tumor or decrease of disease symptoms. Other assays will be known to one of ordinary skill in the art and can be employed for measuring the level of the response.

The doses of sarcoma-associated polypeptide compositions (e.g., polypeptide, peptide, antibody, cell or nucleic acid) administered to a subject can be chosen in accordance with different parameters, in particular in accordance with the mode of administration used and the state of the subject. Other factors include the desired period of treatment. In the event that a response in a subject is insufficient at the initial doses applied, higher doses (or

effectively higher doses by a different, more localized delivery route) may be employed to the extent that patient tolerance permits.

In general, for treatments for eliciting or increasing an immune response, doses of sarcoma-associated antigen are formulated and administered in doses between 1 ng and 1 mg, and preferably between 10 ng and 100 µg, according to any standard procedure in the art. Where nucleic acids encoding sarcoma-associated polypeptides or variants thereof are employed, doses of between 1 ng and 0.1 mg generally will be formulated and administered according to standard procedures. Other protocols for the administration of sarcoma-associated polypeptide compositions will be known to one of ordinary skill in the art, in which the dose amount, schedule of injections, sites of injections, mode of administration (e.g., intra-tumoral) and the like vary from the foregoing. Administration of sarcoma-associated polypeptide compositions to mammals other than humans, e.g. for testing purposes or veterinary therapeutic purposes, is carried out under substantially the same conditions as described above.

Where sarcoma-associated polypeptides are used for vaccination, modes of administration which effectively deliver the sarcoma-associated polypeptide and adjuvant, such that an immune response to the polypeptide is increased, can be used. For administration of a sarcoma-associated polypeptide in adjuvant, preferred methods include intradermal, intravenous, intramuscular and subcutaneous administration. Although these are preferred embòdiments, the invention is not limited by the particular modes of administration disclosed herein. Standard references in the art (e.g., Remington's Pharmaceutical Sciences, 18th edition, 1990) provide modes of administration and formulations for delivery of immunogens with adjuvant or in a non-adjuvant carrier.

The pharmaceutical compositions may contain suitable buffering agents, including: acetic acid in a salt; citric acid in a salt; boric acid in a salt; and phosphoric acid in a salt.

The pharmaceutical compositions also may contain, optionally, suitable preservatives, such as: benzalkonium chloride; chlorobutanol; parabens and thimerosal.

The pharmaceutical compositions may conveniently be presented in unit dosage form and may be prepared by any of the methods well-known in the art of pharmacy. All methods include the step of bringing the active agent into association with a carrier which constitutes one or more accessory ingredients. In general, the compositions are prepared by uniformly and intimately bringing the active compound into association with a liquid carrier, a finely divided solid carrier, or both, and then, if necessary, shaping the product.

Compositions suitable for oral administration may be presented as discrete units, such as capsules, tablets, lozenges, each containing a predetermined amount of the active compound. Other compositions include suspensions in aqueous liquids or non-aqueous liquids such as a syrup, elixir or an emulsion.

Compositions for parenteral administration include sterile aqueous or non-aqueous solutions, suspensions, and emulsions. Examples of non-aqueous solvents are propylene glycol, polyethylene glycol, vegetable oils such as olive oil, and injectable organic esters such as ethyl oleate. Aqueous carriers include water, alcoholic/aqueous solutions, emulsions or suspensions, including saline and buffered media. Parenteral vehicles include sodium chloride solution, Ringer's dextrose, dextrose and sodium chloride, and lactated Ringer's or fixed oils. Intravenous vehicles include fluid and nutrient replenishers, electrolyte replenishers (such as those based on Ringer's dextrose), and the like. Preservatives and other additives may also be present such as, for example, antimicrobials, anti-oxidants, chelating agents, and inert gases, and the like.

The pharmaceutical agents of the invention may be administered alone, in combination with each other, and/or in combination with other anti-cancer drug therapies and/or treatments. These therapies and/or treatments may include, but are not limited to: surgical intervention, chemotherapy, radiotherapy, and adjuvant systemic therapies.

The invention also provides a pharmaceutical kit comprising one or more containers comprising one or more of the pharmaceutical compounds or agents of the invention. Additional materials may be included in any or all kits of the invention, and such materials may include, but are not limited to buffers, water, enzymes, tubes, control molecules, etc. The kit may also include instructions for the use of the one or more pharmaceutical compounds or agents of the invention for the treatment of cancer.

The invention includes kits for assaying the presence of sarcoma-associated antigens and/or antibodies that specifically bind to sarcoma-associated polypeptides. An example of such a kit may include the above-mentioned polypeptides bound to a substrate, for example a dipstick, which is dipped into a blood or body fluid sample of a subject. The surface of the substrate may then be processed using procedures well known to those of skill in the art, to assess whether specific binding occurred between the polypeptides and agents (e.g. antibodies) in the subject's sample. For example, procedures may include, but are not limited to, contact with a secondary antibody, or other method that indicates the presence of specific binding.

Another example of a kit may include an antibody or antigen-binding fragment thereof, that binds specifically to a sarcoma-associated antigen. The antibody or antigen-binding fragment thereof, may be applied to a tissue or cell sample from a patient with cancer and the sample then processed to assess whether specific binding occurs between the antibody and an antigen or other component of the sample. In addition, the antibody or antigen-binding fragment thereof, may be applied to a body fluid sample, such as serum, from a subject, either suspected of having cancer, diagnosed with cancer, or believed to be free of cancer. As will be understood by one of skill in the art, such binding assays may also be performed with a sample or object contacted with an antibody and/or sarcoma-associated antigen that is in solution, for example in a 96-well plate or applied directly to an object surface.

Another example of a kit of the invention is a kit that provides components necessary to determine the level of expression of one or more sarcoma-associated nucleic acid molecules of the invention. Such components may include primers useful for amplification of one or more sarcoma-associated nucleic acid molecules and/or other chemicals for PCR amplification.

Another example of a kit of the invention is a kit that provides components necessary to determine the level of expression of one or more sarcoma-associated nucleic acid molecules of the invention using a method of hybridization.

The foregoing kits can include instructions or other printed material on how to use the various components of the kits for diagnostic purposes.

The invention further includes nucleic acid or protein microarrays (including antibody arrays) for the analysis of expression of sarcoma-associated antigens or nucleic acids encoding such antigens. In this aspect of the invention, standard techniques of microarray technology are utilized to assess expression of the sarcoma-associated antigens and/or identify biological constituents that bind such antigens. The constituents of biological samples include antibodies, lymphocytes (particularly T lymphocytes), and the like. Microarray substrates include but are not limited to glass, silica, aluminosilicates, borosilicates, metal oxides such as alumina and nickel oxide, various clays, nitrocellulose, or nylon. The microarray substrates may be coated with a compound to enhance synthesis of a probe (peptide or nucleic acid) on the substrate. Coupling agents or groups on the substrate can be used to covalently link the first nucleotide or amino acid to the substrate. A variety of coupling agents or groups are known to those of skill in the art. Peptide or nucleic acid

probes thus can be synthesized directly on the substrate in a predetermined grid. Alternatively, peptide or nucleic acid probes can be spotted on the substrate, and in such cases the substrate may be coated with a compound to enhance binding of the probe to the substrate. In these embodiments, presynthesized probes are applied to the substrate in a precise, predetermined volume and grid pattern, preferably utilizing a computer-controlled robot to apply probe to the substrate in a contact-printing manner or in a non-contact manner such as ink jet or piezo-electric delivery. Probes may be covalently linked to the substrate. Nucleic acid probes preferably are linked using UV irradiation or heat.

Protein microarray technology, which is also known by other names including protein chip technology and solid-phase protein array technology, is well known to those of ordinary skill in the art and is based on, but not limited to, obtaining an array of identified peptides or proteins on a fixed substrate, binding target molecules or biological constituents to the peptides, and evaluating such binding. See, e.g., G. MacBeath and S.L. Schreiber, "Printing Proteins as Microarrays for High-Throughput Function Determination," *Science* 289(5485):1760-1763, 2000.

Targets are peptides or proteins and may be natural or synthetic. The tissue may be obtained from a subject or may be grown in culture (e.g. from a cell line).

In some embodiments of the invention, one or more control peptide or protein molecules are attached to the substrate. Preferably, control peptide or protein molecules allow determination of factors such as peptide or protein quality and binding characteristics, reagent quality and effectiveness, hybridization success, and analysis thresholds and success.

Nucleic acid arrays, particularly arrays that bind sarcoma-associated antigens, also can be used for diagnostic applications, such as for identifying subjects that have a condition characterized by aberrant sarcoma-associated antigen expression. Nucleic acid microarray technology, which is also known by other names including: DNA chip technology, gene chip technology, and solid-phase nucleic acid array technology, is well known to those of ordinary skill in the art and is based on, but not limited to, obtaining an array of identified nucleic acid probes on a fixed substrate, labeling target molecules with reporter molecules (e.g., radioactive, chemiluminescent, or fluorescent tags such as fluorescein, Cye3-dUTP, or Cye5-dUTP), hybridizing target nucleic acids to the probes, and evaluating target-probe hybridization. A probe with a nucleic acid sequence that perfectly matches the target sequence will, in general, result in detection of a stronger reporter-molecule signal than will probes with less perfect matches. Many components and techniques utilized in nucleic acid

microarray technology are presented in *The Chipping Forecast, Nature Genetics*, Vol.21, Jan 1999, the entire contents of which is incorporated by reference herein.

According to the invention, probes are selected from the group of nucleic acids including, but not limited to: DNA, genomic DNA, cDNA, and oligonucleotides; and may be natural or synthetic. Oligonucleotide probes preferably are 20 to 25-mer oligonucleotides and DNA/cDNA probes preferably are 500 to 5000 bases in length, although other lengths may be used. Appropriate probe length may be determined by one of ordinary skill in the art by following art-known procedures. In one embodiment, preferred probes are sets of one or more of the sarcoma-associated nucleic acid molecules as described herein. Probes may be purified to remove contaminants using standard methods known to those of ordinary skill in the art such as gel filtration or precipitation.

In one embodiment, the microarray substrate may be coated with a compound to enhance synthesis of the probe on the substrate. Such compounds include, but are not limited to, oligoethylene glycols. In another embodiment, coupling agents or groups on the substrate can be used to covalently link the first nucleotide or oligonucleotide to the substrate. These agents or groups may include, for example, amino, hydroxy, bromo, and carboxy groups. These reactive groups are preferably attached to the substrate through a hydrocarbyl radical such as an alkylene or phenylene divalent radical, one valence position occupied by the chain bonding and the remaining attached to the reactive groups. These hydrocarbyl groups may contain up to about ten carbon atoms, preferably up to about six carbon atoms. Alkylene radicals are usually preferred containing two to four carbon atoms in the principal chain. These and additional details of the process are disclosed, for example, in U.S. Patent 4,458,066, which is incorporated by reference in its entirety.

In one embodiment, nucleic acid probes are synthesized directly on the substrate in a predetermined grid pattern using methods such as light-directed chemical synthesis, photochemical deprotection, or delivery of nucleotide precursors to the substrate and subsequent probe production.

Targets for microarrays are nucleic acids selected from the group, including but not limited to: DNA, genomic DNA, cDNA, RNA, mRNA and may be natural or synthetic. In all embodiments, nucleic acid target molecules from human tissue are preferred. The tissue may be obtained from a subject or may be grown in culture (e.g. from a cell line).

In embodiments of the invention one or more control nucleic acid molecules are attached to the substrate. Preferably, control nucleic acid molecules allow determination of

factors such as nucleic acid quality and binding characteristics, reagent quality and effectiveness, hybridization success, and analysis thresholds and success. Control nucleic acids may include but are not limited to expression products of genes such as housekeeping genes or fragments thereof.

## Examples

### Materials and Methods

***Cell Lines, Tissues, Sera and RNA***

SW1045, SW982, and Fuji synovial sarcoma cell lines were obtained from the cell repository of the Ludwig Institute for Cancer Research (LICR), New York Branch at the Memorial Sloan-Kettering Cancer Center. Tumor tissues and sera were obtained from Memorial Sloan-Kettering Cancer Center, Weill Medical College of Cornell University and Aichi Cancer Center Research Center, Nagoya Japan. Normal tissue RNA preparations were purchased from Clontech laboratories Incorporated (Palo Alto, CA) and Ambion Incorporated (Austin, Texas). Total RNA from tumor tissues was prepared by the guanidinium thiocyanate method.

***SEREX analysis of cDNA expression libraries***

Poly(A)+ RNA from two sarcoma cell lines, SW1045 and SW982, was prepared using the Fast Track mRNA Purification Kit (Invitrogen, Life Technologies, Carlsbad, CA). Poly(A)+ RNA from normal testis was purchased from CLONTECH. Separate cDNA libraries were constructed for each of these in the ZAP Express vector (Stratagene, La Jolla, CA) according to the manufacturer's instructions using 5 μg polyA+ mRNA. Libraries containing 1-2 x $10^6$ primary recombinants were obtained and were not amplified before immunoscreening.

To remove serum antibodies reactive with vector-related antigens, sera was absorbed against *E.coli*/bacteriophage lysates prepared in the following manner. Wild-type lambda ZAP Express bacteriophage at a concentration of 5,000 pfu (plaque-forming units) per 15 cm plate were amplified in *E.coli* XL1 Blue MRF' overnight in 100 ml NZY/0.7% agarose. 10 ml of binding buffer (0.1M NaHCO3, pH 8.3) was then added to the plates, and the plates were gently agitated at 4°C for 15 hours. The resultant supernatants were collected and

residual *E.coli* were lysed by sonication. The lysates were then coupled to CNBr-Sepharose 4B (Amersham Pharmacia Biotech, Piscataway, NJ) according to the manufacturer's instructions. Patient sera (1:10 dilution) were absorbed by batch absorption with an equal volume of Sepharose 4B coupled *E.coli*/phage lysates at 4°C for 6 hours. This procedure was repeated a total of three times and was followed by a 15 hour incubation with nitrocellulose filters precoated with proteins derived from *E. coli* and *E.coli*/phage lysates. Library screenings were performed as previously described (Scanlan, M. J., et al. Characterization of human colon cancer antigens recognized by autologous antibodies. *Int. J. Cancer* 1998; 76: 652-8. Scanlan, M. J., et al. Antigens recognized by autologous antibody in patients with renal-cell carcinoma. *Int. J. Cancer* 1999; 83: 456-64.) A total of five independent SEREX immunoscreenings of the cDNA libraries were undertaken. Sera from 2 sarcoma patients were used independently, at a dilution of 1:200, to immunoscreen the cDNA libraries. A total of 2.5-5.0 x $10^5$ or 1.75 x $10^6$ recombinants were screened per serum/cDNA library combination. Serum reactive phage clones were converted to plasmid forms and subjected to DNA sequencing (Cornell University DNA Services, Ithaca, NY).

***Determination of serum antibody reactivity***

Two assays were used to determine serological reactivity, an ELISA-based method and a bacteriophage expression method. With regard to CT antigens, serum antibody reactivity was determined by ELISA as previously described (Stockert E, et al. 1998. A survey of the humoral immune response of cancer patients to a panel of human tumor antigens. J Exp Med 187:1349-54.) Briefly, recombinant proteins (NY-ESO-1, SSX-2, MAGE-A1, MAGE-A3, MAGE-A4, MAGE-A10, CT7 and CT10) were produced in *E.coli* by transfection with pQE30 expression vectors (Qiagen, Chatsworth, CA) according to the manufacturer's protocol. 10 ng of recombinant protein (1μg/ml) was absorbed to TC microwell plates (Nalge Nunc International Corp., Naperville, IL) for 15 hours at 4°C. After washing with PBS, plates were then blocked with 2% BSA and incubated with diluted (1:100 – 1:25,000) patient sera for 2 hours at room temperature. Following a PBS wash step, 10 μl of a 1:5000 dilution of alkaline phosphatase-conjugated goat anti-human IgG secondary antibody (Southern Biotechnology, Birmingham, AL) was added to each well and incubated for 1 hour at room temperature. Following a PBS wash step, plates were incubated with 10ul/well Attophose substrate (JBL Scientific, San Louis Obispo, CA) for 25 min, and the fluorescence was then read by a Cyto-Fluor 2350 (Millipore, Bedford, MA).

In the case of SEREX-defined sarcoma antigens, a previously described serum antibody detection array (SADA or spot immunoassay (Scanlan MJ, et al. Humoral immunity to human breast cancer: antigen definition and quantitative analysis of mRNA expression. Cancer Immunity 1:4 [epub]; Scanlan MJ. et al. 2002. Cancer-Related Serological Recognition of Human Colon Cancer: Identification of Potential Diagnostic and Immunotherapeutic Targets. Cancer Res. 2002 Jul. 15; 16 (14): 4041-7.) was used to determine serological reactivity.

Preabsorbed serum samples from 39 sarcoma patients and 33 healthy blood donors were evaluated for the presence of IgG antibody reactive to a panel of SEREX-defined sarcoma antigens, identified herein, in the following manner. Precut nitrocellulose membranes (80 X 120mm) were precoated with a layer (approximately 0.2 mm) of growth media (NZY/0.7% Agarose/2.5 mM isopropyl-β-D-thiogalactopyranoside) and placed on a reservoir layer of NZY/0.7% Agarose in a 86 X 128mm Omni Tray (Nalge Nunc). $5.0 \times 10^3$ pfu per µl of bacteriophage encoding individual SEREX-defined tumor antigens were mixed with an equal volume of exponentially growing *E. coli* XL-1 Blue MRF' and spotted (0.7 µl aliquots) on the precoated nitrocellulose membranes using a 96 pin replicator (Nalge Nunc). Membranes were incubated for 15 hours at 37°C and then processed as per the standard SEREX protocol (Scanlan, et al., Int. J. Cancer 1998; 76: 652-8; Scanlan, et al., Int. J. Cancer 1999; 83: 456-64). Briefly, plates were blocked in 0.5% non-fat dried milk; incubated in 10 ml of a 1:200 dilution of sera at room temperature for 15 hours; and then incubated in a 1:3000 dilution of alkaline phosphatase conjugated, Fc fragment specific, goat anti-human IgG (Jackson Immunoresearch laboratories Inc., West Grove PA). Serum IgG reactivity was detected with the alkaline phosphatase substrate, 4-nitro blue tetrazolium chloride/5-bromo-4-chloro-3-indolyl-phosphate.

***Reverse Transcriptase- PCR (RT-PCR) analysis***

The cDNA preparations used as templates in the RT-PCR reactions were prepared using the Superscript first strand synthesis kit (Invitrogen Life Technologies, Carlsbad, CA) according to the manufacturer's instructions using 2.5 µg of total RNA. For evaluation of CT antigens expression in sarcoma cell lines, PCR primers specific for NY-ESO-1, LAGE-1, MAGE-1, MAGE-3, MAGE-4, MAGE-10, SCP-1, BAGE, CT7, SSX-1, SSX-2, and SSX-4 were synthesized commercially (Invitrogen Life Technologies) using published primer

sequences (van der Bruggen P, et al. 1991. A gene encoding an antigen recognized by cytolytic T lymphocytes on a human melanoma. Science 254:1643-47, Gaugler, B., et al. Human gene MAGE-3 codes for an antigen recognized on a melanoma by autologous cytolytic T lymphocytes. *J. Exp. Med.* 1994; 179: 921-30, Chen, Y. -T., et al. A testicular antigen aberrantly expressed in human cancers detected by autologous antibody screening. *Proc. Natl. Acad. Sci. USA.* 1997; 94: 1914-18; Boel, P., et al., and van der Bruggen, P. BAGE: a new gene encoding an antigen recognized on human melanomas by cytolytic T lymphocytes. *Immunity* 1995; *2:* 167-75. (PMID: 7895173); Sahin U, et al. 1998. Expression of multiple cancer/testis antigens in breast cancer and melanoma: basis for polyvalent CT vaccine strategies. Int J Cancer 78:387-89, Lethe B, et al. 1998. LAGE-1, a new gene with tumor specificity. Int. J. Cancer 76:903-8, Türeci Ö, et al. 1998. Expression of SSX genes in human tumors. Int J Cancer 77:19-23, Gure AO, et al. 1997. SSX: a multigene family with several members transcribed in normal testis and human cancer. Int J Cancer 72:965-971). PCR primers specific for SEREX-defined antigens were also synthesized commercially (Invitrogen Life Technologies) and their sequences are as follows: NY-SAR-12 forward, TggCgCAgAAAggAAAAggAAAAT (SEQ ID NO: 91); NY-SAR-12 reverse, AgAggTAgCTggCAggATgTTAg (SEQ ID NO: 92); NY-SAR-35 forward,CTTggTgCgATCAgCCTTAT (SEQ ID NO: 93); NY-SAR-35 reverse,TTgATgCATgAAAACAgAACTC (SEQ ID NO: 94); NY-SAR-41forward, AgAATTggCAgAggCTCgTCATCA (SEQ ID NO: 95); NY-SAR-41 reverse, TTCCAATTTTgCCTTCTCTAACTg (SEQ ID NO: 96); NY-SAR-73 forward, CCCggAgCACgTCgAggTCTAC (SEQ ID NO: 135); NY-SAR-73 reverse, ggTgAggggCCCAggAAgC (SEQ ID NO: 136); NY-SAR-78 forward, CACAATgTATCCTgTTgAAAg (SEQ ID NO: 137); NY-SAR-78 reverse, gAgATgATACATTCTTCCAg (SEQ ID NO: 138); NY-SAR-92 forward, CTTCCgCCAACTCCTCCTACC (SEQ ID NO: 139); NY-SAR-92 reverse, gATgCCCgTgTCTTgTCCTT (SEQ ID NO: 140); NY-SAR-96 forward, CACTAggCTgCTgAggAAgAT (SEQ ID NO: 141); NY-SAR-96 reverse, gTTTTggTgggCAgCATTgAg (SEQ ID NO: 142); NY-SAR-97 forward, ggACCACCCCAAATAgAA (SEQ ID NO: 143); NY-SAR-97 reverse, CCACCAgCTCAggAAgA (SEQ ID NO: 144); NY-SAR-110 forward, TCTgATggAgCggTgggATgC (SEQ ID NO: 145); NY-SAR-110 reverse, gTgTgCCTCggCTTCTTTCTTC (SEQ ID NO: 146).

RT-PCR was performed in the following manner. Twenty-five µl PCR reaction mixtures, consisting of 2 µl cDNA, 0.2 mM dNTP, 1.5 mM $MgCl_2$, 0.25 µM gene specific forward and reverse primers, and 2.5 U Platinum Taq DNA polymerase (Invitrogen Life Technologies), were heated to 94°C for 2 min., followed by 35 thermal cycles of 94°C for 30 seconds, 55°C for 30 seconds and 72°C for 1 min., and a final cycle of 94°C for 30 seconds, 55°C for 30 seconds and 72°C for 5 min. Thermal cycling was performed using a GeneAmp PCR System 9700 (Applied Biosystems, Foster City, CA). Resultant PCR products were analyzed in 2% Agarose/Tris- Acetate-EDTA gels.

***Real-time quantitative reverse transcription (RT)-PCR***

The concentration of NY-SAR-35 mRNA transcripts in normal tissues was measured by real-time RT-PCR using cDNA preparations derived from lung cancer specimens and 16 different normal tissues that had been normalized for 6 housekeeping genes (Clontech). Gene-specific TaqMan probes and PCR primers were designed using Primer Express software (PE Biosystems, Foster City, CA). PCR reactions were prepared using 2.5 µl of cDNA diluted in TaqMan PCR Master Mix (PE Biosystems) supplemented with 200 nM 6-carboxy-fluorescein labeled gene-specific TaqMan probe, and a predetermined, optimum concentration of gene specific forward and reverse primers (300-900 nM). Triplicate PCR reactions were prepared for each cDNA sample. PCR consisted of 40 cycles of 95°C denaturation (15 seconds) and 60°C annealing/extension (60 seconds). Thermal cycling and fluorescent monitoring were performed using an ABI 7700 sequence analyzer (PE Biosystems). The point at which the PCR product is first detected above a fixed threshold, termed cycle threshold (Ct), was determined for each sample. The abundance of gene-specific transcripts in normal tissues was determined by comparison with a standard curve generated from the Ct values of known concentrations of plasmid DNA template encoding NY-SAR-35.

TaqMan primers were as follow: NY-SAR-35 forward, TggTgCgATCAgCCTTATCC (SEQ ID NO: 147); NY-SAR-35 reverse, CggTTCgCTCCTCCAgAA (SEQ ID NO: 148). TaqMan probe: NY-SAR-35, TgTCTgCCCATTTATTgCCgCTCTCT (SEQ ID NO: 149).

***Northern Blot Analysis.***

A Northern blot containing poly A+ RNA (2 μg/lane) from various normal tissues was obtained commercially (Clontech). An NY-SAR-35 cDNA probe (bp 263-1029) was labeled using the Bright Star Psoralen-Biotin Kit (Ambion Inc., Austin, TX) and hybridized to the membrane for 15 hours at 68°C. After washing, the hybridization signal was developed using the Bright Star Bio-Detect Kit, according to the manufacturer's instructions (Ambion).

***Southern Blot analysis***

Genomic DNA was extracted from normal human testis, and samples (10 μg) were independently digested with *EcoRI, HindIII,* and *BamHI* at 37°C overnight. The DNA was then separated on 0.7% agarose gel and blotted onto a nylon transfer membrane. An NY-SAR-35 cDNA probe (bp 252-1029) was radiolabeled with $^{32}$P-dCTP using a random-primer DNA labeling kit (Roche Molecular Biochemicals, Indianapolis, IN). The blot was hybridized to a $^{32}$P labeled probe at 68°C. After 15 hours of hybridization, the membrane was washed under high stringency conditions (0.1XSSC, 0.5%SDS at 60°C) and exposed for autoradiography.

***Example 1 – Results from the first round of immunoscreenings by SEREX analysis***

***Identification of human sarcoma antigens by SEREX analysis***

Preliminary studies were carried out to determine optimum sources of target antigens and immunoreactive patient sera. Three sarcoma cell lines were typed for expression of NY-ESO-1, LAGE-1, MAGE-1, MAGE-3, MAGE-4, MAGE-10, BAGE, SCP-1, CT7, SSX-1, SSX-2, and SSX-4 transcripts by RT-PCR. As shown in Table 2, all 3 sarcoma cell lines expressed at least one of the transcripts in this panel. Specifically, the SW982 and SW1045 synovial sarcoma cell lines expressed 8 and 10 of the 12 CT antigen transcripts in the panel, respectively, while Fuji synovial sarcoma cells expressed 4/12 CT antigen transcripts.

**Table 2. Cancer/Testis antigen expression in sarcoma cell lines**

| CT Antigen | Cell Line | | |
|---|---|---|---|
| | SW982 (synovial) | SW1045 (synovial) | Fuji |
| NY-ESO-1 | + | + | + |
| LAGE-1 | Neg | + | + |
| MAGE-A1 | + | + | Neg |

| | | | |
|---|---|---|---|
| MAGE-A3 | + | + | Neg |
| MAGE-A4 | + | + | |
| MAGE-A10 | + | + | Neg |
| BAGE | + | + | Neg |
| SCP-1 | Neg | Neg | Neg |
| CT7 | + | + | Neg |
| SSX1 | Neg | + | Neg |
| SSX2 | Neg | Neg | |
| SSX4 | + | + | Neg |
| **Totals** | **8/12** | **10/12** | **4/12** |

In order to identify a subset of sarcoma patients that are actively mounting an immune response against tumor antigens, sera from 54 sarcoma patients (various histologies) were tested by ELISA (Stockert E, et al. 1998. A survey of the humoral immune response of cancer patients to a panel of human tumor antigens. J Exp Med 187:1349-54) for the presence of antibodies against a panel of 8 CT antigens consisting of: NY-ESO-1, SSX-2, MAGE-A1, MAGE-A3, MAGE-A4, MAGE-A10, CT7 and CT10. Only 2/54 sarcoma patients, a malignant fibrous histocytoma (MFH) and fibrosarcoma patient (FS), had detectable serum antibodies against a CT antigen, while the remaining 52 patients lacked detectable anti-CT antigen antibodies. Both seropositive patients had antibodies to NY-ESO-1 but lacked antibodies to the other 7 CT antigens tested. Fibrosarcoma tissue from the NY-ESO-1 seropositive patient, FS, was available for CT antigen typing by RT-PCR and was found to express 11/12 different CT antigen transcripts (NY-ESO-1, LAGE-1, MAGE-A1, -A3, -A4, -A10, BAGE, CT7, SSX1, -2 and -4). Tissue from the NY-ESO-1 seropositive patient, MFH, was not available for CT antigen typing by RT-PCR.

Although it was determined that CT antigen expression is frequent in sarcoma tissue, serum antibody responses were not as frequent. This lack of immunogenicity in sarcoma may be an indication of immune escape by sarcoma cells, whereby the immune system fails to recognize CT antigens and eliminate tumor cells expressing these antigens, resulting in the expansion of a homogenous population CT antigen expressing sarcoma cells. Relevant escape mechanisms include defective antigen presentation (Garrido F, Algarra I. MHC antigens and tumor escape from immune surveillance. Adv Cancer Res 2001;83:117-58) and/or production of immuno-inhibitory cytokines, such as TGF-β and IL-10 (Conrad CT, et al. Differential expression of transforming growth factor beta 1 and interleukin 10 in progressing and regressing areas of primary melanoma. J Exp Clin Cancer Res 1999 Jun;18(2):225-32). It is also possible that homogeneous NY-ESO-1 and MAGE expression

in synovial sarcoma (Jungbluth AA, et al. 2001. Monophasic and biphasic synovial sarcomas abundantly express cancer/testis antigen NY-ESO-1 but not MAGE-A1 or CT7. Int J Cancer 94:252-6; Antonescu CR, et al. MAGE antigen expression in monophasic and biphasic synovial sarcoma. Hum Pathol 2002 Feb;33(2):225-9), as opposed to heterogeneous CT antigen expression observed in many other tumor types (Jungbluth AA, et al. 2001. Immunohistochemical analysis of NY-ESO-1 antigen expression in normal and malignant human tissues. Int J Cancer 92:856-60; Jungbluth AA, et al. 2000. Expression of MAGE-antigens in normal tissues and cancer. Int J Cancer 85:460-5), may also be a contributing factor to immune escape.

These 2 patients were chosen as the serum sources for SEREX immunoscreening of cDNA libraries prepared from the SW982 and SW1045 synovial sarcoma cell lines. A total of 4 SEREX immunoscreenings were performed, leading to the identification of 72 distinct sarcoma antigens, designated NY-SAR-1 through NY-SAR-72. As shown in Table 3, immunoscreening with sera from an NY-ESO-1 serum antibody positive MFH patient led to the identification of 28 antigens, including 8 overlapping antigens derived from both the SW982 and SW1045 cDNA libraries, as well as 13 antigens derived solely from the SW982 cDNA library, and 7 antigens derived solely from the SW1045 cDNA library.

Immunoscreening with sera from an NY-ESO-1 serum antibody positive fibrosarcoma patient defined 46 antigens, including 2 overlapping antigens derived from both the SW982 and SW1045 cDNA libraries, as well as 25 antigens derived solely from the SW982 cDNA library, and 19 antigens derived solely from the SW1045 cDNA library. There was little overlap between the antigens recognized by serum antibodies from the MFH and FS patients. Only three antigens, NY-SAR-1/TMF1, NY-SAR-4/FH and NY-SAR-17/LAGE-1 were identified with both the MFH and FS sera. Because serological reactivity to NY-ESO-1 was the criteria used in selecting sera for cDNA library screening, mutual immunoreactivity to the highly homologous (84% amino acid identity) NY-SAR-17/LAGE-1 antigen was expected, and, although not intending to be bound by a particular theory, is likely to be due to shared epitopes. The 72 antigens (Tables 4-6) represent 58 known proteins and 14 uncharacterized gene products.

**Table 3. Immunoscreening of synovial sarcoma cDNA expression libraries with allogeneic sarcoma patient sera**

| Sarcoma Serum | Synovial | Number of | Number of | Total |
|---|---|---|---|---|

| | sarcoma cDNA expression library | recombinants screened | different antigens identified | number of distinct antigens |
|---|---|---|---|---|
| Malignant Fibrous Histocytoma | SW982 | $5 \times 10^5$ | 21 | 28 |
| | SW1045 | $5 \times 10^5$ | 15 | |
| Fibrosarcoma | SW982 | $2.5 \times 10^5$ | 27 | 46 |
| | SW1045 | $2.5 \times 10^5$ | 21 | |

**Table 4. SEREX-defined sarcoma antigens: antigens reactive with sera from multiple cancer patients**

| NY-SAR-Antigen | Identity (Unigene cluster) | Reactivity with Sarcoma Sera | Source of Reactive Sera[1] | SEREX Database ID Number[2] of Equivalent Isolate (Tumor Source[1]) |
|---|---|---|---|---|
| 2 | STAU (Hs.6113) | 2/39 | MFH (#3), OS (#2) | 614 (PRC), 1273 (BC) |
| 4 | FH (Hs.75653) | 5/39 | MFH (#3), OS (#4, #7), ES (#1), FS (#2) | No Match |
| 12 | NESG1 (Hs.158450) | 2/39 | MFH (#3), LS (#4) | No Match |
| 13 | ACTN1 (Hs.119000) | 1/39 | MFH (#3) | 855 (BC) |
| 15 | RBM6 (Hs.173993) | 1/39 | MFH (#3) | 76 (LC) |
| 16 | FLJ12785 (Hs.192742) | 1/39 | MFH (#3) | 756 (TALL) |
| 17 | LAGE-1a (Hs.87225) | 2/39 | MFH (#3), FS (#2) | 1160 (BC) |
| 18 | SSSCA1 (Hs.25723) | 1/39 | MFH (#3) | 1799 (CC) |

| | | | | |
|---|---|---|---|---|
| 28 | MGC:9727 (Hs.11065) | 1/39 | MFH (#3) | 71 (BC) |
| 30 | SNK (Hs.3838) | 2/39 | FS (#2), RS (#1) | No Match |
| 44 | LGALS1 (Hs.227751) | 1/39 | FS (#2) | 704 (RC) |
| 47 | MIF (Hs.73798) | 1/39 | FS (#2) | 989 (MEL) |
| 50 | PYCR1 (Hs.79217) | 3/39 | FS (#2), MFH (#2, #4) | No Match |
| 71 | None (Hs.314941) | 1/39 | FS (#2) | 1938 (GL) |
| 72 | HSPE1 (Hs.1197) | 1/39 | FS (#2) | 882 (HC), 1202 (MEL) |

Antigens did not react with sera from normal blood donors (0/33).[1] Abbreviations: BC, breast cancer; CC, colon cancer; ES, Ewing sarcoma; FS, fibrosarcoma; GC, gastric cancer; GL, glioma; HC, hepatocellular carcinoma; LC, lung cancer; LS, leimyosarcoma; MEL, melanoma; MFH, malignant fibrous histocytoma; OC, ovarian cancer; OS, osteosarcoma; PRC, prostate cancer; RC, renal cancer; RS, rhabdomyosarcoma; TALL, T-cell acute lymphocytic leukemia.[2] SEREX database ID numbers from the LICR's SEREX database (http://www.licr.org/SEREX.html).

**Table 5. SEREX-defined sarcoma antigens: antigens reactive with sera from both normal donors and sarcoma patients**

| NY-SAR-Antigen | Identity (Unigene cluster) | SEREX Database ID Number[1] of Equivalent Isolate (Tumor Source[2]) | Reactivity with Normal Sera | Reactivity with Sarcoma Sera |
|---|---|---|---|---|
| 1 | TMF1 (Hs.267632) | 246 (G), 1241 (BC) | 2/33 | 3/39 |
| 3 | KIAA1536 (Hs.156667) | 89 (BR) | 2/33 | 3/39 |
| 6 | RHAMM (Hs.72550) | 1513 (OC) | 1/33 | 3/39 |
| 7 | PINCH (Hs.112378) | 344 (CC), 550 (GC), 1152 (RC), 1281 (BR) | 16/21 | 14/39 |

| | | | | |
|---|---|---|---|---|
| 10 | KIAA0603 (Hs.173802) | No Match | 11/33 | 4/39 |
| 11 | U2AF1RS2 (Hs.171909) | 430 (RC), 786 (HD), 1236 (BC), 1334 (GC) | 6/33 | 17/39 |
| 14 | SC65 (Hs.207251) | No Match | 8/33 | 4/39 |
| 19 | HEF1 (Hs.80261) | 421 (RC) | 3/33 | 7/39 |
| 22 | NELIN (Hs.216381) | No Match | 4/33 | 19/39 |
| 29 | FLJ13441 (Hs.232146) | 974 (PC) | 6/33 | 3/39 |
| 31 | HUMAUANTIG (Hs.75528) | 1017 (BC), 1331 (GC), 1475 (OC) | 2/33 | 6/39 |
| 32 | PDAP1 (Hs.278426) | No Match | 4/33 | 8/39 |
| 33 | SURF6 (Hs.274430) | No Match | 2/33 | 2/39 |
| 41 | None (Hs.166670) | No Match | 1/33 | 1/39 |
| 45 | STIP1 (Hs.75612) | 430 (RC) | 4/33 | 2/39 |
| 53 | FXYD5 (Hs.333418) | No Match | 1/33 | 1/39 |
| 54 | LMOD1 (Hs.79386) | No Match | 7/33 | 13/39 |
| 55 | RBM10 (Hs.154583) | No Match | 1/33 | 1/39 |
| 58 | LIP8 (Hs.348012) | No Match | 1/33 | 3/39 |
| 61 | ZNF282 (Hs.58167) | No Match | 1/33 | 2/39 |
| 64 | USP16 (Hs.99819) | No Match | 2/33 | 2/39 |
| 65 | FDFT1 (Hs.48876) | No Match | 2/33 | 1/39 |
| 66 | ROCK1 (Hs.109450) | 444 (RC) | 1/33 | 1/39 |
| 68 | P38IP (Hs.333500) | No Match | 1/33 | 3/39 |

[1] The LICR's SEREX database ID numbers from http://www.licr.org/SEREX.html.
[2] Abbreviations: BC, breast cancer; CC, colon cancer; HD, Hodgkins disease; GC, gastric cancer; OC, ovarian cancer; PC, pancreatic cancer; RC, renal cancer.

**Table 6. SEREX-defined sarcoma antigens: antigens reactive with sera from a single sarcoma patient**

| NY-SAR-Antigen | Gene Identity (Unigene Cluster) | NY-SAR-Antigen | Gene Identity (Unigene Cluster) |
|---|---|---|---|
| 5 | TBC1D1(Hs.278586) | 42 | BTG3 (Hs.77311) |
| 8 | BIRC2 (Hs.289107) | 43 | SSX4 (Hs.278632) |
| 9 | ATP5B (Hs.25) | 46 | ARNTL2 (Hs.222024) |
| 20 | TCEB3 (Hs.155202) | 48 | MGC20533 (Hs.69280) |
| 21 | GTF3C3 (Hs.90847) | 49 | EMK1 (Hs.157199) |
| 23 | C20orf81 (Hs.29341) | 51 | EDF1 (Hs.174050) |
| 24 | None (not clustered) | 52 | Actin (Hs.288061) |
| 25 | PDE4DIP (Hs.265848) | 56 | MLF1Hs.85195) |
| 26 | PIASX-BETA (Hs.111323) | 57 | GCN5L2 (Hs.101067) |
| 27 | FLJ10330(Hs.342307) | 59 | UPF3B (Hs.103832) |
| 34 | SEC23B (Hs.173497) | 60 | EGLN1 (Hs.6523) |
| 35 | None (Hs.128580) | 62 | AD034(Hs.281397) |
| 36 | SSX1(Hs.194759) | 63 | USP19(Hs.301373) |
| 37 | MP1 (Hs.260116) | 67 | LUC7L (Hs.16803) |
| 38 | HMG20B (Hs.32317) | 69 | ARL1 (Hs.242894) |
| 39 | PSMD4 (Hs.148495) | 70 | RPL10A (Hs.334895) |
| 40 | INPP1 (Hs.32309) | | |

[1]Antigens reacted with sera from single sarcoma patient (1/39), but not with sera from normal individuals (0/33). The antigens listed had no matches with existing entries in the SEREX database (http://www.licr.org/SEREX.html).

The nucleotide sequences of all uncharacterized gene products (NY-SAR-3, -10, -16, -22, -23, -24, -27, -28, -29, -35, -41, -48, - 62, -71) have been deposited in the GenBank database (SEQ ID NOs: 1-14, respectively). The cDNA sequences encoding the 72 sarcoma antigens were also compared to sequences deposited in the SEREX database accessible through a website of the Ludwig Institute for Cancer Research (http://www.licr.org/SEREX.html). Examination of this database revealed that 21 of the 72 sarcoma antigens defined in this study (29%) were also identified through SEREX analysis of other tumor types (Tables 4 and 5).

***Reactivity patterns of sera from normal individuals and cancer patients with SEREX-defined sarcoma antigens***

To determine whether immune recognition of the isolated antigens was cancer-related, allogeneic sera samples obtained from 33 normal blood donors and 39 sarcoma patients (various histologies) were tested for reactivity against the 72 sarcoma antigens defined in the current study using serum antibody detection arrays (SADA). Twenty-four of the 72 antigens (33%) had a serological profile that was not restricted to cancer patients, as evidenced by their reactivity with normal sera. These antigens have been listed in Table 5.

Sera from two normal individuals and three sarcoma patients reacted with NY-SAR-1/TMF1, suggesting the reactivity was unrelated to cancers. With one notable exception (NY-SAR-22/ NELIN), the frequency of antibody responses to 23 of the 24 antigens associated with normal sera reactivity was similar in normal blood donors and cancer patients. In the case of NY-SAR-22/ NELIN (UniGene cluster Hs.216381), the frequency of antibody responses was considerably higher in cancer patients, in which 19/39 (49%) of sarcoma patients and 4/33 (12%) of normal individuals had a detectable antibody response. The remaining 48 antigens had a cancer-related serological profile, reacting only with sera from cancer patients.

The 48 antigens having a cancer-related serological profile could be subdivided into 4 categories; a) antigens identified by serum from only a single sarcoma patient; b) antigens that reacted with sera from a single sarcoma patient and, as determined by an analysis of the SEREX database, with sera from patients having other forms of cancer; c) antigens that reacted exclusively with sera from 2 or more sarcoma patients; and d) antigens that reacted with sera from 2 or more sarcoma patients and with sera from patients having other forms of

cancer. Of the 48 antigens having a cancer-related serological profile, 33 antigens reacted with sera from a single sarcoma patient (Table 6).

As shown in Table 4, the remaining 15 antigens reacted with sera from 2 or more cancer patients, but not with sera from normal individuals. Nine antigens reacted with sera from a single sarcoma patient, and with sera from patients with other tumor types (NY-SAR-13, -15, -16, -18, -28, -44, -47, -71, -72). Four antigens reacted exclusively with sera from 2 or more sarcoma patients (NY-SAR, -4, -12, -30, -50). The remaining two antigens, NY-SAR-2/STAU and the CT antigen, NY-SAR-17/ LAGE-1A, reacted with sera from 2 or more sarcoma patients and with sera from patients with other types of cancer. A cancer-related serological response to NY-SAR-4/FH occurred most frequently. In this case, serum samples from 5/39 (13%) sarcoma patients were reactive with NY-SAR-4/FH, including 2/10 sera samples from osteosarcoma patients, 1/6 sera samples from malignant fibrous histiocytoma patients, 1/2 patients sera samples from fibrosarcoma patients, and 1/7 sera samples from Ewing sarcoma patients. No serological responses to NY-SAR-4/FH were detected in normal blood donors.

This serological response to NY-SAR-4/FH is of interest as germ-line mutations in the FH gene have been associated with a predisposition to uterine and cutaneous leiomyomata and also renal cell carcinoma (Tomlinson IP, et al. Germline mutations in FH predispose to dominantly inherited uterine fibroids, skin leiomyomata and papillary renal cell cancer. Nat Genet 2002 Apr;30(4):406-10) and is a target of somatic mutation in sarcoma (Kiuru, M., et al. (2002) Cancer Res. 62, 4554-4557) suggesting that the immune response is directed against mutated epitopes.

***Expression patterns of mRNA encoding serologically defined sarcoma antigens in normal and malignant tissues***

A preliminary *in silico* mRNA expression profile of all gene products identified in this study was carried out based on the tissue distribution of expressed sequence tags (ESTs) in the human EST database. Products with no EST matches, or those having EST matches limited to tumor tissue, fetal tissue, and/or less than 3 normal adult tissues were further examined by RT-PCR. Gene products with restricted EST profiles include the three well-characterized cancer-testis antigens, LAGE-1/NY-SAR-17, NY-SAR-36/SSX1, and NY-SAR-43/SSX4, which are expressed exclusively in normal testis and a range of different tumor types (Lethe B, et al. 1998. LAGE-1, a new gene with tumor specificity. Int. J. Cancer

76:903-8; Türeci Ö, et al. 1998. Expression of SSX genes in human tumors. Int. J. Cancer 77:19-23; Gure AO, et al. 1997. SSX: a multigene family with several members transcribed in normal testis and human cancer. Int. J. Cancer 72:965-971), and 3 putative tissue restricted antigens, including a known gene product, nasopharyngeal specific protein 1(NESG1)/NY-SAR-12 (Li Z, Yao K, Cao Y. Molecular cloning of a novel tissue-specific gene from human nasopharyngeal epithelium. Gene 1999 Sep 3;237(1):235-40), and 2 uncharacterized gene products, NY-SAR-35 (UniGene cluster Hs.128580) and NY-SAR-41 (UniGene cluster Hs.166670). With the exception of serum reactivity to NY-SAR-41 occurring in 1/33 normal blood donors, these differentially expressed antigens showed a cancer-related serological profile.

As shown in Figure 1A, mRNA expression patterns of NY-SAR-12, -35, and -41 were examined in 17 different human tissues by RT-PCR. NESG1/NY-SAR-12 mRNA was detected in normal placenta, testis, colon, lung, and ovary (0/12 other normal tissues). NY-SAR-35 mRNA was detected only in normal testis (0/15 other normal tissues), while a lower molecular weight transcript was detected in normal ovary. NY-SAR-41 was detected in normal testis, fetal brain, colon, lung, and bladder (0/12 other normal tissues). As shown in Figure 1B, the testis restricted expression pattern of NY-SAR-35 was confirmed by real time quantitative RT-PCR at 40 amplification cycles. In these studies, NY-SAR-35 was expressed in normal testis at a level corresponding to 83.2 ag, which was more than 1000 times the level detected in the remaining 15 normal tissues.

The expression of NY-SAR-35 mRNA was also examined in 26 sarcoma specimens of various histologies, and was detected in fibrosarcoma and rhabdomyosarcoma specimens (2/26), as well as the SW1045 synovial sarcoma cell line (Table 7 and Figure 1C). With regard to other tumor types, transcripts encoding NY-SAR-35 were detected in 1/16 (6%) melanoma specimens, 5/29 (21%) lung cancer specimens, and 3/13 (23%) breast cancer specimens. NY-SAR-35 mRNA was not detected in small number of colon cancer specimens (0/9) or in small numbers of renal cancer specimens (0/8). Thus, on the basis of its immunogenicity in cancer patients, and its restricted mRNA expression profile, NY-SAR-35 can be considered a novel CT antigen.

**Table 7. Expression of NY-SAR-35 in sarcoma, sarcoma cell lines and other malignant tissues**

| Histology | Expression Frequency |
|---|---|
| **Sarcomas** | |
| Synovial sarcoma | 0/8 |
| Leimyosarcoma | 0/4 |
| Malignant Fibrous Histocytoma | 0/4 |
| Ewing Sarcoma | 0/2 |
| Osteosarcoma | 0/2 |
| Rhabdomyosarcoma | 1/1 |
| Fibrosarcoma | 1/1 |
| Liposarcoma | 0/1 |
| Neurosarcoma | 0/1 |
| Chondrosarcoma | 0/1 |
| DFSP | 0/1 |
| SW1045 synovial sarcoma cell line | positive |
| SW982 synovial sarcoma cell line | negative |
| Fuji synovial sarcoma cell line | negative |
| **Other Malignancies** | |
| Melanoma | 1/16 |
| Lung Cancer | 5/29 |
| Colon Cancer | 0/9 |
| Breast Cancer | 3/13 |
| Renal Cancer | 0/8 |
| Esophageal Cancer | 1/12 |
| Ovarian Cancer | 1/12 |
| Gastric Cancer | 5/6 |

***The NY-SAR-35 gene, transcript and putative protein and orthologous gene***

An analysis of the human genome database, mapped the NY-SAR-35 cDNA sequence to Xq28, approximately 5.9Mbp downstream (3') of the CT10/MAGE-E1 gene and 6.8 Mbp

upstream (5') of the NY-ESO-1 gene. The NY-SAR-35 gene is approximately 44 kb in length and spans 6 exons. Analyses of the human genome databases (NCBI GenBank, http://www.ncbi.nlm.nih.gov/genome, and Celera Genomics, Rockville, MD, www.celera.com) revealed no genomic sequences of high similarity, suggesting that it is a single copy gene with no additional family members. These results were verified by probing Southern blots of human genomic DNA with the NY-SAR-35 cDNA.

The present SEREX immunoscreening provided 4 overlapping NY-SAR-35 cDNA clones, ranging from 677-767 bp in length, all contained identical 3' sequences originating from the poly A region. The NY-SAR-35 cDNA sequence was identical to 3 ESTs (GenBank accession nos. AA909915, AA906131, and AW593050) which were all derived from the NFL_T_GBC_S1 mixed tissue (fetal lung, testis, germinal center B cell) cDNA library and found in UniGene cluster Hs.128580 as well as 4 ESTs (GenBank accession nos. BC034320, AK098602, BG771667 and BI465380) derived from a testis cell line and found in Unigene cluster Hs.375082. As shown in Figure 1D, Northern blot analysis revealed a single NY-SAR-35 mRNA transcript of 1.1 kB in normal testis, indicating the SEREX-defined clones and EST sequences represent partial transcripts. To obtain a full-length NY-SAR-35 transcript, 5' RACE was performed, yielding 262 bp of additional 5' DNA sequence. Thus, the total length of the NY-SAR-35 transcript is 1029 bp (SEQ ID NO: 10, GenBank accession no. AY211917), a size that is in agreement with the 1.1 kb hybridization signal seen in Northern blots of testis mRNA probed with NY-SAR-35 cDNA.

The NY-SAR-35 transcript encodes an open reading frame of 255 amino acids (SEQ ID NO: 55, bp 68-895) with a predicted molecular mass of 29.2kDa. It is identical to a hypothetical protein, XM098959, predicted from Genefinder analysis of human chromosome X sequences. The putative NY-SAR-35 protein has a signal peptide, a transmembrane domain and a cysteine-rich trefoil/P-domain, found in several secreted proteins of the gastrointestinal tract (Hoffmann W, Hauser F. The P-domain or trefoil motif: a role in renewal and pathology of mucous epithelia? Trends Biochem Sci 1993 Jul;18(7):239-43). These data suggest that NY-SAR-35 is a-secreted or membrane bound protein.

To identify a murine orthologue of NY-SAR-35, the putative human NY-SAR-35 protein sequence was used to search a translated nonredundant nucleotide database by using the TBLASTN tool of the NCBI (www.ncbi.nlm.nih.gov/blast/Blast.cgi). A hypothetical mouse protein, termed XP_150408, generated from a conceptual translation of the mouse X chromosome, was found to have 57% identity (49/85 amino acids) with NY-SAR-35. Using

nucleotide primers corresponding to sequences encoding XP_150408, 5' and 3' RACE reactions were undertaken by using mouse testis cDNA. By combining 5' and 3' RACE products, a 1,202 bp cDNA was identified (GenBank accession no. AY214130, SEQ ID NO: 133). This cDNA encoded a putative full length mouse protein of 238 amino acids (SEQ ID NO: 134) which is 41% identical to human NY-SAR-35, with conservation of the trefoil and transmembrane domains. This murine NY-SAR-35 (mNY-SAR-35) cDNA sequence was used to search mouse genome sequences (www.ncbi.nlm.nih.gov/genome/seq/MmBlast.html) yielding an identical genome sequence, NW 042622, from mouse chromosome X. Analysis of this sequence showed the mNY-SAR-35 gene is composed of approximately 42,600 nucleotides and seven exons.

***Example 2 –Analysis of the NY-SAR-35 protein and its expression***

***Purification of recombinant NY-SAR-35 protein in E. coli to produce monoclonal antibodies and to perform ELISA assays***

There are four ATG codons in exon 1 of the NY-SAR-35 gene. It is expected that the fourth ATG codon in the full length sequence of NY-SAR-35 is the first ATG codon of the translated NY-SAR-35 sequence. It appears then that the predicted protein has two interesting domains. The protein revealed two distinctive hydrophobic domains followed by two hydrophilic turns. One hydrophobic domain is a signal peptide, which are predicted in proteins with cleavage sites between amino acids 25 and 26 with SignalP software tool available at the website http://www.cbs.dtu.dk/services/SignalP. The other hydrophobic region is predicted to be a transmembrane domain with the TMHMM2.0 program available at the website http://www.cbs.dtu.dk/services/TMHMM/TMHMM2.0b.guide.html. Therefore, the NY-SAR-35 gene encodes a signal peptide and a transmembrane domain (Figure 2).

Three kinds of NY-SAR-35 vectors were designed for the purification of the proteins in an *E. coli* expression system (pET System) (Novagen, Madison, WI). The first encoded the largest possible NY-SAR-35 protein from the first ATG codon (SEQ ID NO: 150), the second encoded the NY-SAR-35 protein from the fourth ATG codon (MH7) (SEQ ID NO: 152), and the third encoded the expected extracellular domain from the fourth ATG codon (SEQ ID NO: 154). An illustration of these vectors is provided below. The expected sizes of the resulting proteins are 29 kD (263 amino acids) (SEQ ID NO: 151), 22 kD (201 amino acids) (SEQ ID NO: 153) and 14.6 kD (133 amino acids) (SEQ ID NO: 155), respectively.

I. Whole protein(NY-SAR-35 from the first ATG)
Vector:pET23a(NdeI/XhoI): C-terminal His tag vector
Primer;
SAR35/NdeI : CACACACA**CATATG**TCTTCACATAGGAGGAAAGCGAAG (SEQ ID NO: 156)
SAR35/XhoI : CACACA**CTCGAG**CTCGTCACCATGTTCCTCACGTC (SEQ ID NO: 157)

**CATATG**TCTTCACATAGGAGGAAAGCGAAGGGGAGGAATAGGAGAAGTCACCG
TGCCATGCGTGTGGCTCACTTAGAGCTGGCAACTTATGAGTTGGCGGCAACTGAG
TCGAATCCCGAGAGCAGCCATCCTGGATACGAGGCCGCCATGGCTGACAGGCCT
CAGCCAGGATGGCGGGAATCTCTAAAGATGCGGGTCAGCAAACCCTTTGGGATG
CTCATGCTCTCCATTTGGATCCTGCTGTTCGTGTGCTACTACCTGTCCTACTACCT
GTGCTCCGGGTCCTCATATTTTGTGCTTGCAAATGGACATATCCTGCCCAACAGT
GAAAATGCTCATGGCCAATCTCTGGAAGAAGATTCCGCATTGGAAGCTTTGCTGA
ATTTTTTCTTTCCAACAACTTGCAATCTGAGGGAAAATCAGGTGGCAAAGCCTTG
TAATGAGCTGCAAGATCTTAGTGAGAGTGAATGTTTGAGACACAAATGCTGTTTT
TCATCATCGGGGACCACGAGCTTCAAATGTTTTGCTCCATTTAGAGATGTGCCTA
AACAGATGATGCAAATGTTTGGGCTTGGTGCGATCAGCCTTATCCTGGTATGTCT
GCCCATTTATTGCCGCTCTCTTTTCTGGAGGAGCGAACCGGCCGATGATTTACAA
AGGCAGGACAACAGAGTTGTAACGGGTTTGAAGAAACAAAGAAGGAAGCGAAA
GAGGAAGTCTGAAATGTTACAGAAAGCAGCAAGAGGACGTGAGGAACATGGTG
ACGAG**CTCGAG**CACCACCACCACCACCACTGA (SEQ ID NO: 150)

MSSHRRKAKGRNRRSHRAMRVAHLELATYELAATESNPESSHPGYEAAMADRPQP
GWRESLKMRVSKPFGMLMLSIWILLFVCYYLSYYLCSGSSYFVLANGHILPNSENAH
GQSLEEDSALEALLNFFFPTTCNLRENQVAKPCNELQDLSESECLRHKCCFSSSGTTSF
KCFAPFRDVPKQMMQMFGLGAISLILVCLPIYCRSLFWRSEPADDLQRQDNRVVTGL
KKQRRKRKRKSEMLQKAARGREEHGDELEHHHHHH (SEQ ID NO: 151)

II. Partial protein(MH7 from the fourth ATG)
Vector; pET23a(NdeI/XhoI)
Primer;
MH7/NdeI: CACACACA**CATATG**CGGGTCAGCAAACCCTTTGGGA (SEQ ID NO: 158)
SAR35/XhoI : CACACA**CTCGAG**CTCGTCACCATGTTCCTCACGTC (SEQ ID NO: 159)

**CATATG**CGGGTCAGCAAACCCTTTGGGATGCTCATGCTCTCCATTTGGATCCTGC
TGTTCGTGTGCTACTACCTGTCCTACTACCTGTGCTCCGGGTCCTCATATTTTGTG
CTTGCAAATGGACATATCCTGCCCAACAGTGAAAATGCTCATGGCCAATCTCTGG
AAGAAGATTCCGCATTGGAAGCTTTGCTGAATTTTTTCTTTCCAACAACTTGCAAT
CTGAGGGAAAATCAGGTGGCAAAGCCTTGTAATGAGCTGCAAGATCTTAGTGAG
AGTGAATGTTTGAGACACAAATGCTGTTTTTCATCATCGGGGACCACGAGCTTCA
AATGTTTTGCTCCATTTAGAGATGTGCCTAAACAGATGATGCAAATGTTTGGGCT
TGGTGCGATCAGCCTTATCCTGGTATGTCTGCCCATTTATTGCCGCTCTCTTTTCT
GGAGGAGCGAACCGGCCGATGATTTACAAAGGCAGGACAACAGAGTTGTAACG
GGTTTGAAGAAACAAAGAAGGAAGCGAAAGAGGAAGTCTGAAATGTTACAGAA

AGCAGCAAGAGGACGTGAGGAACATGGTGACGAG**CTCGAG**CACCACCACCACC ACCACTGA (SEQ ID NO: 152)

MRVSKPFGMLMLSIWILLFVCYYLSYYLCSGSSYFVLANGHILPNSENAHGQSLEEDS ALEALLNFFFPTTCNLRENQVAKPCNELQDLSESECLRHKCCFSSSGTTSFKCFAPFR DVPKQMMQMFGLGAISLILVCLPIYCRSLFWRSEPADDLQRQDNRVVTGLKKQRRK RKRKSEMLQKAARGREEHGDELEHHHHHH (SEQ ID NO: 153)

III. Expected extracellular domain of NY-SAR-35 from the fourth ATG

Vector:pET23a(NdeI/XhoI)

Primer;

MH7/NdeI: CACACACA**CATATG**CGGGTCAGCAAACCCTTTGGGA (SEQ ID NO: 160)

MH7/XhoI: CACACA**CTCGAG**CATTTGCATCATCTGTTTAGGC (SEQ ID NO: 161)

**CATATG**CGGGTCAGCAAACCCTTTGGGATGCTCATGCTCTCCATTTGGATCCTGC TGTTCGTGTGCTACTACCTGTCCTACTACCTGTGCTCCGGGTCCTCATATTTTGTG CTTGCAAATGGACATATCCTGCCCAACAGTGAAAATGCTCATGGCCAATCTCTGG AAGAAGATTCCGCATTGGAAGCTTTGCTGAATTTTTTCTTTCCAACAACTTGCAAT CTGAGGGAAAATCAGGTGGCAAAGCCTTGTAATGAGCTGCAAGATCTTAGTGAG AGTGAATGTTTGAGACACAAATGCTGTTTTTCATCATCGGGGACCACGAGCTTCA AATGTTTTGCTCCATTTAGAGATGTGCCTAAACAGATGATGCAAATG**CTCGAG**C ACCACCACCACCACCACTGA (SEQ ID NO: 154)

MRVSKPFGMLMLSIWILLFVCYYLSYYLCSGSSYFVLANGHILPNSENAHGQSLEEDS ALEALLNFFFPTTCNLRENQVAKPCNELQDLSESECLRHKCCFSSSGTTSFKCFAPFR DVPKQMMQMLEHHHHHH (SEQ ID NO: 155)

Protein expression was induced in *E. coli*. Three colonies of each domain cloned plasmid were selected and cultured by IPTG induction for 4 hours. When total proteins were separated by SDS-electrophoresis and stained by Simply Blue SafeStain (Invitrogen) the highly expressed protein bands were not detected. However, when total proteins, separated by SDS-polyacrylamide gel, were immunoblotted using an anti-His epitope antibody, the His-tagged NY-SAR-35 proteins were detected. The results are shown in Figure 3 with the expected sizes of the expressed proteins.

***Functional study of NY-SAR-35***

Most cancer-testis antigens (11/13) have been found to be expressed in non-malignant human kidney embryonic 293 cell while NY-SAR-35 is not. Human 293 cell and monkey Cos-1 cells were used to stably express the NY-SAR-35 gene for functional and immunolocalization studies.

The expected NY-SAR-35 open reading frame (including the 5' untranslated region) was cloned into pcDNA3.1/V5-HisA vector which had a C-terminal fusion tag (V5 epitope and 6XHis epitope). The cloned NY-SAR-35 nucleotide sequence and expected amino acid sequence are as follows:

**A. Cloned NY-SAR-35 nucleotide sequence**

GAATTCCTTCTGGGCCACGGACTGCCGGACCGTTGGGCTGTGAGGCAGCGTCTCAGCGAGGC
**EcoRI**
GGCACCCGGAGCCATGTCTTCACATAGGAGGAAAGCGAAGGGGAGGAATAGGAGAAGTCAC

CGTGCCATGCGTGTGGCTCACTTAGAGCTGGCAACTTATGAGTTGGCGGCAACTGAGTCGAAT

CCCGAGAGCAGCCATCCTGGATACGAGGCCGCCATGGCTGACAGGCCTCAGCCAGGATGGCG

GGAATCTCTAAAGATGCGGGTCAGCAAACCCTTTGGGATGCTCATGCTCTCCATTTGGATCCT

GCTGTTCGTGTGCTACTACCTGTCCTACTACCTGTGCTCCGGGTCCTCATATTTTGTGCTTGCA

AATGGACATATCCTGCCCAACAGTGAAAATGCTCATGGCCAATCTCTGGAAGAAGATTCCGCA

TTGGAAGCTTTGCTGAATTTTTTCTTTCCAACAACTTGCAATCTGAGGGAAAATCAGGTGGCA

AAGCCTTGTAATGAGCTGCAAGATCTTAGTGAGAGTGAATGTTTGAGACACAAATGCTGTTTT

TCATCATCGGGGACCACGAGCTTCAAATGTTTTGCTCCATTTAGAGATGTGCCTAAACAGATG

ATGCAAATGTTTGGGCTTGGTGCGATCAGCCTTATCCTGGTATGTCTGCCCATTTATTGCCGCT

CTCTTTTCTGGAGGAGCGAACCGGCCGATGATTTACAAAGGCAGGACAACAGAGTTGTAACG

GGTTTGAAGAAACAAAGAAGGAAGCGAAAGAGGAAGTCTGAAATGTTACAGAAAGCAGCAA

GAGGACGTGAGGAACATGGTGACGAG**CTCGAG**TCTAGAGGGCCCTTCGAA***GGTAAGCCTAT***
**XhoI** **V5 epitope**
***CCCTAACCCTCTCCTCGGTCTCGATTCTACG***CGTACCGGT***CATCATCACCATCACCAT***TGA
**His tag**

(SEQ ID NO: 162)

**B. Expected amino acid sequence and expected size of expressed proteins**

→ 31Kd → 29Kd
EFLLGHGLPDRWAVRQRLSEAAPGAMSSHRRKAKGRNRRSHRAMRVAHLELATYELAATESNPE
→ 26Kd → 24Kd
SSHPGYEAAMADRPQPGWRESLKMRVSKPFGMLMLSIWILLFVCYYLSYYLCSGSSYFVLANGHI

LPNSENAHGQSLEEDSALEALLNFFFPTTCNLRENQVAKPCNELQDLSESECLRHKCCFSSSGTTSF

KCFAPFRDVPKQMMQMFGLGAISLILVCLPIYCRSLFWRSEPADDLQRQDNRVVTGLKKQRRKRK

RKSEMLQKAARGREEHGDELESRGPFE**GKPIPNPLLGLDSTRTGHHHHHH**
**V5 epitope** **His tag**

(SEQ ID NO: 163)

Human 293 cells and monkey Cos-1 cell that stably express the NY-SAR-35 gene were tested by RT-PCR and Western blotting. The cells transfected with 0.5 μg

pcDNA3.1/V5/HisA/NY-SAR-35 plasmid were selected with 1mg/ml neomycin for 14 days. Clones were picked and expanded for an additional 1 month and analyzed for NY-SAR-35 mRNA and protein expression. Three sets of NY-SAR-35 5'/3' primers were used and are provided below:

lane 1 (ORF including 5' untranslated region)
GGGAATTCATGTCTTCACATAGGAGGAAAGCG/CACACACTCGAGCTCGTCACCA TGTTCCTCACGTC (SEQ ID NO: 164)

lane2( ORF from the first ATG)
CACACACACATATGTCTTCACATAGGAGGAAAGCGAAG/CACACACTCGAGCTCG TCACCATGTTCCTCACGTC (SEQ ID NO: 165)

lane 3(ORF from the fourth ATG)
CACACACACATATGCGGGTCAGCAAACCCTTTGGGA/CACACACACATATGTCTT CACATAGGAGGAAAGCGAAG (SEQ ID NO: 166)

lane 4 (p53 5'/3') TACTCCCCTGCCCTCAACAAG/CTCAGGCGGCTCATAGGG (SEQ ID NO: 167)

Whole cell extracts were made from the same cloned cells. Total proteins were separated by SDS-polyacrylamide gel electrophoresis and immunoblotted to detect NY-SAR-35 proteins by anti-V5 epitope monoclonal antibody(Invitrogen). The size of the stably expressed NY-SAR-35 proteins in 293 and Cos-1 cells was found to be 24kD. This is, therefore, consistent with translation of NY-SAR-35 beginning at the fourth ATG.

***Example 3 – Results from the second round of immunoscreenings by SEREX analysis***

***Identification of human sarcoma antigens by SEREX analysis***

Serum from the two NY-ESO-1 seropositive patients (FS and MFH) were again used to immunoscreen cDNA libraries prepared from the SW982 and SW1045 synovial sarcoma cell lines, both of which were shown to express eight or more known CT antigen transcripts (Table 2). Sera from the FS patient was also used to immunoscreen a cDNA library derived from normal testis. In total, the results from Examples 1 and 3 represent five independent SEREX immunoscreenings performed, which lead to the identification of 113 distinct antigens, designated NY-SAR-1 through NY-SAR-113.

The 113 SEREX-defined antigens represent 91 known proteins and 22 uncharacterized gene products (novel, ESTs, KIAA series, FLJ series, ORFs, DKFZ series).

In addition to the uncharacterized gene products described above in Example 1 (NY-SAR-3, -10, -16, -22, -23, -24, -27, -29, -35, -41, -48 and -71) additional immunoscreening identified another 11 uncharacterized gene products (NY-SAR-77, -79, -80, -84, -88, -91, -95, -97, -104, -105 and -113). All of the sequences for these uncharacterized gene products have been deposited in the GenBank database and given the sequential accession numbers AY211909-AY211931. In terms of the serum sources, 27 of the 113 antigens were identified by using sera from a MFH patient and 86 were identified with FS sera. Of the 113 antigens identified, 95 were unique to a particular cDNA library screening and 18 antigens were identified in more than one library. This underlines the beneficial nature of incorporating multiple cDNA libraries into large-scale SEREX analyses of the cancer immunome.

***Seroepidemiology of SEREX-defined sarcoma antigens***

The cDNA sequences encoding the 113 sarcoma antigens were compared with sequences deposited in the cancer immunome or SEREX database (http://www2.licr.org/CancerImmunomeDB, formerly www.licr.org/SEREX.html). These comparisons are in addition to the comparisons presented above in Example 1. In a preliminary analysis, it was found that 39 of the 113 sarcoma antigens defined in this study (34%) were also identified through SEREX analysis of other tumor types (Table 8). Table 9 below provides a complete list of all 113 antigens along with their respective Unigene cluster information, if any. These results represent the information available after all rounds of immunoscreening. Contrary to the results shown, NY-SAR-39, -57, -61, -63 and -64 after the first round of immunoscreenings had not been found in the SEREX database.

**Table 8. Immunomic analysis of sarcoma/testis antigens: Reactivity with sera from sarcoma patients, patients with other forms of cancer, and normal individuals**

| NY-SAR-antigen | Gene identity (ugene cluster) | Cancer patient seroreactivity* | Normal seroreactivity |
|---|---|---|---|
| 1 | TMF1 (Hs.267632) | GC, BC, CC, SRC | 2/33 |
| 2 | STAU (Hs.6113) | PC, BC, SRC | 3/30 |
| 3 | KIAA1536 (Hs.156667) | BC, SRC | 2/33 |
| 6 | RHAMM (Hs.72550) | OC, SRC | 1/33 |
| 7 | PINCH (Hs.112378) | CC, GC, RC, BC HN, ESO, AML, SRC | 16/21 |
| 11 | U2AF1RS2 (Hs.171909) | RC, HD, BC, GC, SRC | 6/33 |
| 13 | ACTN1 (Hs.119000) | BC, SRC | 5/30 |
| 15 | RBM6 (Hs.173993) | LC, SRC | 0/33 |

| | | | |
|---|---|---|---|
| 16 | FLJ12785 (Hs.192742) | TALL, SRC | 0/33 |
| 17 | LAGE-1a (Hs.87225) | BC, SRC | 0/33 |
| 18 | SSSCA1(Hs.25723) | CC, SRC | 0/33 |
| 19 | HEF1 (Hs.80261) | RC, SRC | 3/33 |
| 28 | PPIL4 (Hs.11065) | BC, SRC | 0/33 |
| 29 | FLJ13441 (Hs.232146) | PN, SRC | 6/33 |
| 31 | AUANTIG (Hs.75528) | BC, GC, OC, SRC | 2/33 |
| 39 | PSMD4 (Hs.148495) | MEL, SRC | 0/33 |
| 44 | LGALS1 (Hs.227751) | RC, SRC | 0/33 |
| 45 | STIP1 (Hs.75612) | RC, SRC | 4/33 |
| 47 | MIF (Hs.73798) | MEL, SRC | 0/33 |
| 57 | GCN5L2 (Hs.101067) | PC, SRC | 0/33 |
| 61 | ZNF282 (Hs.58167) | RC, SRC | 1/33 |
| 63 | USP19 (Hs.301373) | OC, SRC | 0/33 |
| 64 | USP16 (Hs.99819) | PN, SRC | 2/33 |
| 66 | ROCK1 (Hs.17820) | RC, BC, CC, SRC | 1/33 |
| 74 | RANBP2 (Hs.199179) | BC, GL, BC, SRC | 2/33 |
| 77 | KIAA0992 (Hs.194431) | PC, SRC | 4/15 |
| 80 | FLJ12577 (Hs.87159) | GC, SRC | 0/33 |
| 81 | SDS3 (Hs.20104) | GC, SRC | 4/16 |
| 82 | NYCO45 (Hs.160881) | CC, SRC | 0/33 |
| 89 | SSX2 (Hs.289105) | BC, MEL, SRC | 0/33 |
| 90 | UACA (Hs.49753) | BC, ESO, SRC | 4/25 |
| 93 | NYBR15 (Hs.178175) | BC, SRC | 1/12 |
| 98 | OIP2 (Hs.274170) | BC, SRC | 0/33 |
| 99 | SSX3 (Hs.178749) | BC, MEL, SRC | 2/30 |
| 101 | RANBP2L1 (Hs.179825) | GL, BC, SRC | 3/33 |
| 102 | RBPJK (Hs.356806) | GC, RC, BC, MEL, SRC | 1/16 |
| 103 | Hsp40 (Hs.94) | HN, NCC, SRC | 0/33 |
| 108 | EIF4G (Hs.25732) | GC, SRC | 5/27 |
| 112 | PMSCL1 (Hs.91728) | CC, SRC | 0/33 |

AML, acute myelogenous leukemia; BC, breast cancer; CC, colon cancer; GC, gastric cancer; GL, glioma; HCC, hepatocellular carcinoma; HN, head and nech cancer; LC, lung cancer; MEL, melanoma; OC, ovarian cancer; PC, prostate cancer; PN, pancreatic cancer; RC, renal cancer; SRC, sarcoma; TALL, T cell acute lymphocytic leukemia. *Determined by sequence comparisons with the SEREX database (www2.licr.org/CancerImmunomeDB/).

**Table 9. Sarcoma/testes antigens defined by serological analysis of cDNA expression libraries**

| NY-SAR-antigen | Gene identity (Unigene Cluster) | Sera source | Library source | NY-SAR-antigen | Gene identity (Unigene Cluster) | Sera source | Library source |
|---|---|---|---|---|---|---|---|
| 1 | TMF1 (Hs.267632) | MFH, FS | A, T | 58 | LIP8 (Hs.348012) | FS | A |
| 2 | STAU (Hs.6113) | MFH | A | 59 | UPF3B (Hs.103832) | FS | A |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 3 | KIAA1536 (Hs.156667) | MFH | A | 60 | EGLN1 (Hs.6523) | FS | A |
| 4 | FH (Hs.75653) | MFH, FS | A, B | 61 | ZNF282 (Hs.58167) | FS | A |
| 5 | TBC1D1(Hs.278586) | MFH | A | 62 | AD034(Hs.281397) | FS | A |
| 6 | RHAMM (Hs.72550) | MFH | A, B | 63 | USP19(Hs.301373) | FS | A |
| 7 | PINCH (Hs.112378) | MFH | A, B | 64 | USP16 (Hs.99819) | FS | B, T |
| 8 | BIRC2 (Hs.289107) | MFH | A, B | 65 | FDFT1 (Hs.48876) | FS | B |
| 9 | ATP5B (Hs.25) | MFH | A, B | 66 | ROCK1 (Hs.17820) | FS | B, T |
| 10 | KIAA0603 (Hs.173802) | MFH | A | 67 | LUC7L (Hs.16803) | FS | B |
| 11 | U2AF1RS2 (Hs.171909) | MFH | A, B | 68 | P38IP (Hs.333500) | FS | B |
| 12 | NESG1 (Hs.158450) | MFH | B | 69 | ARL1 (Hs.242894) | FS | B |
| 13 | ACTN1 (Hs.119000) | MFH | A | 70 | RPL10A (Hs.334895) | FS | B |
| 14 | SC65 (Hs.207251) | MFH | A | 71 | EST (Hs.314941) | FS | B |
| 15 | RBM6 (Hs.173993) | MFH | A | 72 | HSPE1 (Hs.1197) | FS | B, T |
| 16 | FLJ12785 (Hs.192742) | MFH | A | 73 | PRM2 (Hs.2324) | FS | T |
| 17 | LAGE-1a (Hs.87225) | MFH, FS | B | 74 | RANBP2 (Hs.199179) | FS | T |
| 18 | SSSCA1 (Hs.25723) | MFH | A, B | 75 | GKAP42 (Hs.36752) | FS | T |
| 19 | HEF1 (Hs.80261) | MFH | A, B | 76 | TIAL1 (Hs.182741) | FS | T |
| 20 | TCEB3 (Hs.155202) | MFH | B | 77 | KIAA0992 (Hs.194431) | FS | T |
| 21 | GTF3C3 (Hs.90847) | MFH | A | 78 | TSP-NY (Hs.97643) | FS | T |
| 22 | NELIN (Hs.216381) | MFH | A | 79 | Novel (not clustered) | FS | T |
| 23 | C20orf81 (Hs.29341) | MFH | A | 80 | FLJ12577 (Hs.87159) | FS | T |
| 24 | None (not clustered) | MFH | A | 81 | SDS3 (Hs.20104) | FS | T |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 25 | PDE4DIP (Hs.265848) | MFH | B | 82 | NYCO45 (Hs.160881) | FS | T |
| 26 | PIASX-BETA (Hs.111323) | MFH | B | 83 | SOX6 (Hs.326876) | FS | T |
| 27 | FLJ10330(Hs.342307) | MFH | B | 84 | DKFZp434 (Hs.131834) | FS | T |
| 28 | PPIL4 (Hs.11065) | FS | B | 85 | RAD50 (Hs.41587) | FS | T |
| 29 | FLJ13441 (Hs.232146) | FS | A | 86 | EPIM (Hs.99865) | FS | T |
| 30 | SNK (Hs.3838) | FS | A | 87 | SOX5 (Hs.87224) | FS | T |
| 31 | HUMAUANTIG (Hs.75528) | FS | A, B, T | 88 | DKFZp564 (Hs.93589) | FS | T |
| 32 | PDAP1 (Hs.278426) | FS | A | 89 | SSX2 (Hs.289105) | FS | T |
| 33 | SURF6 (Hs.274430) | FS | B | 90 | UACA (Hs.49753) | FS | T |
| 34 | SEC23B (Hs.173497) | FS | B | 91 | FLJ11730 (Hs.17118) | FS | T |
| 35 | EST (Hs.128580) | FS | B, T | 92 | ESTs (Hs.368781) | FS | T |
| 36 | SSX1(Hs.194759) | FS | B, T | 93 | NYBR15 (Hs.178175) | FS | T |
| 37 | MP1 (Hs.260116) | FS | A, T | 94 | CG005 (Hs.23518) | FS | T |
| 38 | HMG20B (Hs.32317) | FS | A | 95 | FLJ10637 (Hs.22595) | FS | T |
| 39 | PSMD4 (Hs.148495) | FS | A | 96 | MCSP (Hs.111850) | FS | T |
| 40 | INPP1 (Hs.32309) | FS | A | 97 | EST (Hs.128836) | FS | T |
| 41 | EST (Hs.166670) | FS | B | 98 | OIP2 (Hs.274170) | FS | T |
| 42 | BTG3 (Hs.77311) | FS | B, T | 99 | SSX3 (Hs.178749) | FS | T |
| 43 | SSX4 (Hs.278632) | FS | B | 100 | PGAM2 (Hs.46039) | FS | T |
| 44 | LGALS1 (Hs.227751) | FS | B | 101 | RANBP2L1 (Hs.179825) | FS | T |
| 45 | STIP1 (Hs.75612) | FS | A | 102 | RBPJK (Hs.356806) | FS | T |
| 46 | ARNTL2 (Hs.222024) | FS | B | 103 | Hsp40 (Hs.94) | FS | T |
| 47 | MIF (Hs.73798) | FS | A | 104 | DKFZp434 (Hs.131834) | FS | T |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 48 | MGC20533 (Hs.69280) | FS | A | 105 | C11orf14 (Hs.32017) | FS | T |
| 49 | EMK1 (Hs.157199) | FS | A | 106 | CEP11 (Hs.97437) | FS | T |
| 50 | PYCR1 (Hs.79217) | FS | A | 107 | UBE1 (Hs.2055) | FS | T |
| 51 | EDF1 (Hs.174050) | FS | A | 108 | EIF4G (Hs.25732) | FS | T |
| 52 | Actin (Hs.288061) | FS | A | 109 | SYNJ1 (Hs.127416) | FS | T |
| 53 | FXYD5 (Hs.333418) | FS | A | 110 | NYD-SP14 (Hs.98105) | FS | T |
| 54 | LMOD1 (Hs.79386) | FS | A | 111 | NDP52 (Hs.154230) | FS | T |
| 55 | RBM10 (Hs.154583) | FS | A | 112 | PMSCL1 (Hs.91728) | FS | T |
| 56 | MLF1(Hs.85195) | FS | A, T | 113 | KIAA0442 (Hs.32168) | FS | T |
| 57 | GCN5L2 (Hs.101067) | FS | A | | | | |

To determine whether immune recognition of these 39 antigens was cancer-related, serum samples from normal individuals (n=33) were tested for reactivity to these antigens. 23 of the 39 antigens (59%) had a serological profile that was not restricted to cancer patients, whereas the remaining 16 antigens had a cancer-related serological profile, reacting only with sera from cancer patients (sarcoma patients and serum source of SEREX database entry), and not with sera from normal individuals. 14 of these 16 antigens reacted only with sera from a single sarcoma patient when tested for reactivity with additional allogeneic sarcoma sera (n=39). The remaining 2 antigens, NY-SAR-17/LAGE-1 and NY-SAR-80/FLJ12577, reacted with 2 of 39 and 3 of 39 sarcoma sera, respectively, and not with sera from normal individuals (n=33).

NY-SAR-80/FLJ12577 is an uncharacterized member of the Mo25 protein family, an evolutionary conserved family of proteins with no known function. Analysis of the tissue distribution and frequency of EST sequences homologous to NY-SAR-80/FLJ12577 indicate widespread mRNA expression, with a preponderance of malignant tissue-derived homologous ESTs suggesting possible overexpression in cancer.

Overall, the relative infrequency of overlapping humoral immune responses among the population of sarcoma patients analyzed is contrary to previous findings for colon (Scanlan MJ. et al. 2002. Cancer-Related Serological Recognition of Human Colon Cancer:

Identification of Potential Diagnostic and Immunotherapeutic Targets. Cancer Res., 2002; Jul. 15; 62(14), 4041-7.), breast (Scanlan MJ, et al. Humoral immunity to human breast cancer: antigen definition and quantitative analysis of mRNA expression. Cancer Immunity 1:4 [epub]) and renal cancers (Scanlan, M. J., et al., and Old, L. J. Antigens recognized by autologous antibody in patients with renal-cell carcinoma. *Int. J. Cancer* 1999; 83: 456-64) in which a subset of antigens were mutually seroreactive in a cancer related manner. These results suggest that the immune response to sarcoma is either highly variable or that distinct sarcoma histiotypes have distinct immunomes.

***Expression patterns of mRNA encoding serologically defined sarcoma/testis antigens in normal and malignant tissues***

In addition to the three well-known CT antigens described in Example 1, NY-SAR-89/SSX-2 and NY-SAR-99/SSX-3 were found to have restricted EST profiles, being expressed exclusively in normal testis and a range of different tumor types (Lethe B, et al. 1998. LAGE-1, a new gene with tumor specificity. Int. J. Cancer 76:903-8; Türeci Ö, et al. 1998. Expression of SSX genes in human tumors. Int J Cancer 77:19-23; Gure AO, et al. 1997. SSX: a multigene family with several members transcribed in normal testis and human cancer. Int J Cancer 72:965-971). Six other putative tissue-restricted antigens were identified, including four other known gene products, NY-SAR-73/Protamine 2 (PRM2, Domenjoud, L., Fronia, C., Uhde, F. & Engel, W. (1998) Nucleic Acids Res. 16, 7773), NY-SAR-78/TSP-NY (UniGene cluster Hs.97643), NY-SAR-96/mitochondrial capsule selenoprotein (MCSP, Aho, H., et al. (1996) Geonomics 32, 184-190) and NY-SAR-110/NYD-SP14 (Hs.98105) and two additional uncharacterized gene products, NY-SAR-92 (Hs.368781) and NY-SAR-97 (not clustered).

Two of the six putative tissue restricted antigens, NY-SAR-73/PRM2 and NY-SAR-110/NYD-SP14, were ubiquitously expressed in a panel of 20 normal tissues as determined by RT-PCR (Table 10). The remaining four genes, in addition to NY-SAR-12/nasopharyngeal specific protein 1 (NESG1, Li Z, Yao K, Cao Y. Molecular cloning of a novel tissue-specific gene from human nasopharyngeal epithelium. Gene 1999 Sep 3;237(1):235-40), NY-SAR-35 and NY-SAR-41, were found to be expressed with frequencies ranging from 1 to 9 of 20 normal tissues. NY-SAR-35 and NY-SAR-78 were both testis-specific. The mRNA expression profiles of NY-SAR-35 and NY-SAR-78 were

then analyzed in various malignant tissues by RT-PCR. Transcripts encoding NY-SAR-78/TSP-NY were not detected in cancer. The tumor specimens examined included, lung cancer (0 of 9), colon cancer (0 of 9), breast cancer (0 of 18), renal cancer (0 of 11), esophageal cancer (0 of 12), ovarian cancer (0 of 14), melanoma (0 of 18) and sarcoma (0 of 8). Thus, although NY-SAR-78/TSP-NY is a "virtual CT antigen" with 100% identity with ESTs derived from prostate cancer and leukemia, its expression in cancer could not be verified in our RT-PCR series.

**Table 10. Analysis of mRNA expression by RT-PCR of 9 of the 113 sarcoma/testis antigens**

| | NY-SAR antigen* | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| **Tissue** | **12** | **35** | **41** | **73** | **78** | **92** | **96** | **97** | **110** |
| Brain | - | - | - | + | - | - | - | - | + |
| Kidney | - | - | - | + | - | - | - | - | + |
| Liver | - | - | - | + | - | - | - | - | + |
| Pancreas | - | - | - | + | - | - | - | - | + |
| Placenta | + | - | - | + | - | - | - | - | + |
| Testis | + | + | + | + | + | + | + | + | + |
| Fetal brain | - | - | + | + | - | - | + | + | + |
| Small intestine | - | - | - | + | - | - | - | - | + |
| Heart | - | - | - | + | - | - | - | - | + |
| Prostate | - | - | - | + | - | - | + | + | + |
| Adrenal | - | - | - | + | - | - | + | + | + |
| Spleen | + | - | - | + | - | + | + | + | + |
| Colon | + | - | + | + | - | - | - | - | + |
| Stomach | - | - | - | + | - | - | - | - | + |
| Lung | + | - | + | + | - | - | - | + | + |
| Bladder | - | - | + | + | - | - | - | + | + |
| Ovary | + | - | + | + | - | - | - | + | + |
| Breast | - | - | - | + | - | - | - | + | + |
| Cervix | - | - | - | + | - | - | - | - | + |
| Skeletal muscle | - | - | - | + | - | - | - | - | + |
| **Total no. of positive tissues** | **6/20** | **1/20** | **6/20** | **20/20** | **1/20** | **2/20** | **5/20** | **9/20** | **20/20** |

*Unigene clusters: NY-SAR-12, Hs.158450; NY-SAR-35, Hs.128580; NY-SAR-41, Hs.166670; NY-SAR-73, Hs.2324; NY-SAR-78, Hs.97643; NY-SAR-92, Hs.368781; NY-SAR-96, Hs.111850; NY-SAR-97, Hs.128836; NY-SAR-110, Hs.98105.

The antigens presented herein are of interest for their immunotherapeutic and diagnostic potential. For example, the six known testis-restricted gene antigens (NY-SAR-17/LAGE-1, NY-SAR-36/SSX1, NY-SAR-43/SSX4, NY-SAR-78/TSP-NY, NY-SAR-

89/SSX2 and NY-SAR-99/SSX3), four novel gene products that are also differentially expressed antigens (NY-SAR-35, -41, -92 and -91) and two tissue-restricted antigens (NY-SAR-12/NESG1 and NY-SAR-96/MCSP) not previously studied in relation to cancer have are potential vaccine targets and/or targets for therapeutic antibodies as well as for diagnosis of cancer, particularly by screening patient samples for antibodies that recognize the proteins.

NY-SAR-35 mRNA was detected in a variety of tumor specimens, such as melanoma (1 of 16 specimens), sarcoma (2 of 26 specimens), lung cancer (5 of 29 specimens), breast cancer (3 of 13 specimens), bladder cancer (5 of 12 specimens), esophageal cancer (1 of 12 specimens) and ovarian cancer (1 of 12 specimens). As also shown before in Example 1, NY-SAR-35 was not detected in colon cancer (n=9) or renal cancer (n=8). The CT-restricted expression profile of NY-SAR-35 was confirmed by real-time quantitative RT-PCR at 40 amplification cycles (Figure 4). In two of the nine non-small lung cancer specimens tested, NY-SAR-35 was expressed at levels that were 0.13 and 0.15 times the level detected in normal testis. In conformity with the proposed nomenclature for CT antigens (Chen YT, et al. 1998. Identification of multiple cancer/testis antigens by allogeneic antibody screening of a melanoma cell line library. Proc. Natl. Acad. Sci. U S A. 95:6919-23), NY-SAR-35 is designated CT-20.

References

1. van der Bruggen P, et al. 1991. A gene encoding an antigen recognized by cytolytic T lymphocytes on a human melanoma. Science 254:1643-47.
2. Gaugler, B., Van den Eynde, B., van der Bruggen, P., Romero, P., Gaforio, J. J., De Plaen, E., Lethe, B., Brasseur, F., and Boon, T. Human gene MAGE-3 codes for an antigen recognized on a melanoma by autologous cytolytic T lymphocytes. *J. Exp. Med.* 1994; 179: 921-30.
3. Kawakami, Y., Eliyahu, S., Delgado, C. H., Robbins, P. F., Rivoltini, L., Topalian, S., L., Miki, T., and Rosenberg, S. A. Cloning of the gene for a shared human melanoma antigen recognized by autologous T cells infiltrating into tumor. *Proc. Natl. Acad. Sci. USA*. 1994; 91: 3515-19.
4. Chen, Y. -T., Scanlan, M. J., Sahin, U., Tureci, O., Gure, A. O., Tsang, S., Williamson, B., Stockert, E., Pfreundschuh, M., and Old, L. J. A testicular antigen aberrantly expressed in human cancers detected by autologous antibody screening. *Proc. Natl. Acad. Sci. USA*. 1997; 94: 1914-18.

5. Jager D, Stockert E, Gure AO, Scanlan MJ, Karbach J, Jager E, Knuth A, Old LJ, Chen YT. Identification of a tissue-specific putative transcription factor in breast tissue by serological screening of a breast cancer library. Cancer Res 2001 Mar 1;61(5):2055-61.

6. Boel, P., Wildmann, C., Sensi, M. L., Brasseur, R., Renauld, J. C., Coulie, P., Boon, T., and van der Bruggen, P. BAGE: a new gene encoding an antigen recognized on human melanomas by cytolytic T lymphocytes. *Immunity* 1995; *2:* 167-75. (PMID: 7895173).

7. Van den Eynde, B., Peeters, O., De Backer, O., Gaugler, B., Lucas, S., and Boon, T. A new family of genes coding for an antigen recognized by autologous cytolytic T lymphocytes on a human melanoma. *J. Exp. Med.* 1995; 182: 689-98. (PMID: 7544395).

8. Skipper JC, Hendrickson RC, Gulden PH, Brichard V, Van Pel A, Chen Y, Shabanowitz J, Wolfel T, Slingluff CL Jr, Boon T, Hunt DF, Engelhard VH. An HLA-A2-restricted tyrosinase antigen on melanoma cells results from posttranslational modification and suggests a novel pathway for processing of membrane proteins. J Exp Med 1996 Feb 1;183(2):527-34.

9. Cox AL, Skipper J, Chen Y, Henderson RA, Darrow TL, Shabanowitz J, Engelhard VH, Hunt DF, Slingluff CL Jr. Identification of a peptide recognized by five melanoma-specific human cytotoxic T cell lines. Science 1994 Apr 29;264(5159):716-9.

10. Pascolo S, Schirle M, Guckel B, Dumrese T, Stumm S, Kayser S, Moris A, Wallwiener D, Rammensee HG, Stevanovic S. A MAGE-A1 HLA-A A*0201 epitope identified by mass spectrometry. Cancer Res 2001 May 15;61(10):4072-7.

11. Sahin, U., Türeci, Ö., Schmitt, H., Cochlovius, B., Johannes, T., Schmits, R., Stenner, F., Luo, G., Schobert, I., and Pfreundschuh, M. Human neoplasms elicit multiple specific immune responses in the autologous host. *Proc. Natl. Acad. Sci. USA* 1995; 92: 11810-13.

12. Scanlan, M. J., Chen, Y. T., Williamson, B., Gure, A. O., Stockert, E., Gordan, J. D., Tureci, O., Sahin, U., Pfreundschuh, M., and Old, L. J. Characterization of human colon cancer antigens recognized by autologous antibodies. *Int. J. Cancer* 1998; 76: 652-8.

13. Scanlan, M. J., Gordan, J. D., Williamson, B., Stockert, E., Bander, N. H., Jongeneel, V., Gure, A. O., Jäger, D., Jäger, E., Knuth, A., Chen, Y.-T., and Old, L. J. Antigens recognized by autologous antibody in patients with renal-cell carcinoma. *Int. J. Cancer* 1999; 83: 456-64.

14. Scanlan MJ, et al. Humoral immunity to human breast cancer: antigen definition and quantitative analysis of mRNA expression. Cancer Immunity 1:4 [epub].

15. Jager D, Stockert E, Jager E, Gure AO, Scanlan MJ, Knuth A, Old LJ, Chen YT. Serological cloning of a melanocyte rab guanosine 5'-triphosphate-binding protein and a chromosome condensation protein from a melanoma complementary DNA library. Cancer Res 2000 Jul 1;60(13):3584-91.

16. Chen, Y. -T., Scanlan, M. J., Obata, Y., and Old, L. J. Identification of human tumor antigens by serological expression cloning. *In:* S. A. Rosenberg (ed.). Principles and Practice of Biologic Therapy of Cancer, pp. 557- 570. Philadelphia: Lippincott Williams & Wilkins, 2000.

17. Sahin U, et al. 1998. Expression of multiple cancer/testis antigens in breast cancer and melanoma: basis for polyvalent CT vaccine strategies. Int J Cancer 78:387-89.

18. Scanlan MJ et al. 2000. Expression of cancer-testis antigens in lung cancer: definition of bromodomain testis-specific gene (BRDT) as a new CT gene, CT9. Cancer Lett. 150:155-64.

19. Van den Eynde BJ van der Bruggen P. 1997. T cell defined tumor antigens. Curr Opin Immunol 9:684-693.

20. Jungbluth AA, et al. 2001. Monophasic and biphasic synovial sarcomas abundantly express cancer/testis antigen NY-ESO-1 but not MAGE-A1 or CT7. Int J Cancer 94:252-6.

21. Antonescu CR, Busam KJ, Iversen K, Kolb D, Coplan K, Spagnoli GC, Ladanyi M, Old LJ, Jungbluth AA. MAGE antigen expression in monophasic and biphasic synovial sarcoma. Hum Pathol 2002 Feb;33(2):225-9.

22. Stockert E, et al. 1998. A survey of the humoral immune response of cancer patients to a panel of human tumor antigens. J Exp Med 187:1349-54.

23. Scanlan MJ. et al. 2002. Cancer-Related Serological Recognition of Human Colon Cancer: Identification of Potential Diagnostic and Immunotherapeutic Targets. Cancer Res., 2002; Jul. 15; 62(14), 4041-7.

24. Lethe B, et al. 1998. LAGE-1, a new gene with tumor specificity. Int. J. Cancer 76:903-8.

25. Türeci Ö, et al. 1998. Expression of SSX genes in human tumors. Int J Cancer 77:19-23.

26. Gure AO, et al. 1997. SSX: a multigene family with several members transcribed in normal testis and human cancer. Int J Cancer 72:965-971.

27. Li Z, Yao K, Cao Y. Molecular cloning of a novel tissue-specific gene from human nasopharyngeal epithelium. Gene 1999 Sep 3;237(1):235-40.

28. Hoffmann W, Hauser F. The P-domain or trefoil motif: a role in renewal and pathology of mucous epithelia? Trends Biochem Sci 1993 Jul;18(7):239-43.

29. Chen YT, et al. 1998. Identification of multiple cancer/testis antigens by allogeneic antibody screening of a melanoma cell line library. Proc. Natl. Acad. Sci. U S A. 95:6919-23.

30. Niwa M, Maruyama H, Fujimoto T, Dohi K, Maruyama IN. Affinity selection of cDNA libraries by lambda phage surface display. Gene 2000 Oct 3;256(1-2):229-36.

31. Le Naour F, Misek DE, Krause MC, Deneux L, Giordano TJ, Scholl S, Hanash SM. Proteomics-based identification of RS/DJ-1 as a novel circulating tumor antigen in breast cancer. Clin Cancer Res 2001 Nov;7(11):3328-35.

32. Yang XF, Wu CJ, McLaughlin S, Chillemi A, Wang KS, Canning C, Alyea EP, Kantoff P, Soiffer RJ, Dranoff G, Ritz J. CML66, a broadly immunogenic tumor antigen, elicits a humoral immune response associated with remission of chronic myelogenous leukemia. Proc Natl Acad Sci U S A 2001 Jun 19;98(13):7492-7.

33. Jungbluth AA, et al. 2001. Immunohistochemical analysis of NY-ESO-1 antigen expression in normal and malignant human tissues. Int J Cancer 92:856-60.

34. Jungbluth AA, et al. 2000. Expression of MAGE-antigens in normal tissues and cancer. Int J Cancer 85:460-5.

35. Garrido F, Algarra I. MHC antigens and tumor escape from immune surveillance. Adv Cancer Res 2001;83:117-58.

36. Conrad CT, Ernst NR, Dummer W, Brocker EB, Becker JC. Differential expression of transforming growth factor beta 1 and interleukin 10 in progressing and regressing areas of primary melanoma. J Exp Clin Cancer Res 1999 Jun;18(2):225-32.

37. Tomlinson IP, et al. Germline mutations in FH predispose to dominantly inherited uterine fibroids, skin leiomyomata and papillary renal cell cancer. Nat Genet 2002 Apr;30(4):406-10.

38. Ayyoub M, et al. 2002. Proteasome-assisted identification of a SSX-2-derived epitope recognized by tumor-reactive CTL infiltrating metastatic melanoma. J Immunol 168:1717-22.

39. Van Der Bruggen, P., Zhang, Y., Chaux, P., Stroobant, V., Panichelli, C., Schultz, E.S., Chapiro, J., Van Den Eynde, B.J., Brasseur, F. & Boon, T. (2002) Immunol. Rev. 188, 51-64.

40. Rosenberg, S.A. (2001) Nature 411, 380-384.

41. Preuss, K.D., Zwick, C., Bormann, C., Neumann, F. & Pfeundschuh, M. (2002) Immunol. Rev. 188, 43-50.

42. Scanlan, M.J., Gure, A.O., Jungbluth, A.A., Old, L.J. & Chen, Y.T. (2002) Immunol. Rev. 188, 22-32.

43. Guillaudeux, T., Gomez, E., Onno, M., Drenou, B., Segretain, D., Alberti, S., Lejeune, H., Fauchet, R., Jegou, B. & Le Bouteiller, P. (1996) Biol. Reprod. 55, 99-110.

44. Domenjoud, L., Fronia, C., Uhde, F. & Engel, W. (1998) Nucleic Acids Res. 16, 7773.

45. Aho, H., Schwemmer, M., Tessman, D., Murphy, D., Mattei, G., Engel, W. & Adham, I.M. (1996) Geonomics 32, 184-190.

46. Naka, N., Araki, N., Nakanishi, H., Itoh, K., Mano, M., Ishiguro, S., de Bruijn, D.R. Myoui, A., Ueda, T. & Yoshikawa, H. (2002) Int. J. Cancer 98, 640-642.

47. Martelange, V., De Smet, C., De Plaen, E., Lurquin, C. & Boon, T. (2000) Cancer Res. 60, 3848-3855.

48. Kiuru, M., Lehtonen, R., Arola, J., Salovaara, R., Jarvinen, H., Aittomaki, K., Sjobeg, J., Visakorpi, T., Knuutila, S., et al. (2002) Cancer Res. 62, 4554-4557.

## EQUIVALENTS

Those skilled in the art will recognize, or be able to ascertain using no more than routine experimentation, many equivalents to the specific embodiments of the invention described herein. Such equivalents are intended to be encompassed by the following claims.

All references disclosed herein are incorporated by reference in their entirety.

What is claimed is: